

Deanship of Graduate Studies

Al- Quds University



**Amplicon-based Next generation sequencing for
Identification of Sand fly and leishmania parasites**

Mohammad Hashem Hosen Altarade

M.Sc. Thesis

Jerusalem- Palestine

1440 \ 2019

Amplicon-based Next generation sequencing for Identification of Sand fly and leishmania parasites

Prepared By:

Mohammad Hashem Hosen Altarade

B.SC in Medical Laboratory science

Al-Quds University

Palestine

Supervisor:

Prof. Ziad Abdeen

Dr. Abed Nasereddin

A thesis Submitted in Partial fulfillment of requirements for
the degree of Master of Biochemistry and Molecular
Biology from Biochemistry and Molecular Biology
Department Deanship of Graduate Studies/ Al- Quds
University

1440 \ 2019

Al Quds University

Deanship of Graduate Studies

Biochemistry and Molecular Biology Department



Thesis Approval

**Amplicon-based Next generation sequencing for
Identification of Sand fly and leishmania parasites**

Prepared By: Mohammad Hashem Hosen Altarade

Registration No: 21512256

Supervisor: Prof. Ziad Abdeen

Co-supervisor: Dr. Abed Nasereddin

Master thesis submitted and accepted, Date: 18/5/2019

The names and signatures of the examining committee members are as follows.

1-Head of Committee: Prof. Ziad Abdeen

signature

2- Co-Supervisor: Dr. Abed Nasereddin

signature

3-Internal Examiner: Dr. Ibrahim Abbasi

signature

4-External Examiner: Dr. Robin Abu Ghazaleh

signature

Jerusalem- Palestine

1440 \ 2019

Dedication

To my mother and father...

To my brothers and sisters...

To my friends...

To all beloved people who supported, assisted and encouraged me.

Mohammad Hashem Hosen Altarade

Declaration:

I certify that this thesis submitted for the degree of Master, is the results of my own research, except where otherwise acknowledged, and that this study (or any part of the same) has not been submitted for a higher degree to any other university of institution.

Signed:

A handwritten signature in blue ink, consisting of a long horizontal stroke followed by a loop and a small flourish.

(Mohammad Hashem Hosen Altarade)

Date: 18/5/2019

Acknowledgements

All thanks be to Allah

I would like to thank my thesis adviser Dr. Abed Nasereddin for his supervision directions and assistance discussion and kindness throughout this study.

My thanks also go to all those who help me at Alabraj Research lab especially for Dr. Suheir Ereqat and Ahmad Abdelkader. Especially thanks to Mr. Samer Sawalhah, Palestinian entomology department Ministry of health. I would like also to acknowledge Al-Quds University for allowing me to get a master degree in Biochemistry and Molecular biology.

Also, I would like to thank my family and my friends for their support.

Mohammad Hashem Hosen Altarade

Table of Contents

Dedication.....	I
Declaration:	i
Acknowledgements	ii
List of abbreviations.....	vii
Abstract	viii
Chapter One: Introduction.....	1
1.1. Leishmaniasis	1
1.2. Leishmaniasis worldwide	2
1.3. Leishmaniasis in Palestine.....	3
1.4. Leishmaniasis treatment.	4
1.5. Leishmania Parasite	5
1.6. Parasite Life cycle	6
1.7. Sand Fly.....	7
1.8. Molecular technology	9
1. 9. Study Objectives.....	9
Chapter Two: Materials and Methods	10
2.1. Sand fly sample collection.....	10
2.2 DNA Extraction.....	10
2.3 Design of NGS specific primers	11
2.4 Amplicon PCR Amplification	13
2.5.1 Leishmania primers optimization	13
2.5.2 Sand fly primer Optimization	15
2.5.3 PCR products purification and NGS	15
2.6 Multiplex PCR and Sand fly primers as limit factor	16
2.7 Selection of Leishmania species specific probes.....	18
2.8 Optimizing the quality copy number for Leishmania.....	19
2.9 Analysis of field Sand flies:	20
2.10 Bioinformatics analysis	22
Workflow design:	22
Sand fly virtual Probes	23
2.11. Statistical analysis and multiple sequence alignment.....	24

Chapter Three: Results	25
3.1 PCRs optimization.....	25
3.2 Optimization with different global worldwide <i>Leishmania</i> species	30
3.3 Species specific probes design.....	31
3.4 Detection of minimal Leishmania ITS-PCR copy number	33
3.5 Field Samples	34
3.6 NGS data and statistical analysis	36
Chapter Four: Discussion	40
Study limitation	44
Recommendations	44
REFERENCES.....	45
.5 الملخص 61	

List of Figures

Figure 1.1: The epidemiology of leishmaniasis diseases worldwide.	3
Figure 1.2: Geographical distribution of leishmaniasis disease in the Palestinian West Bank region between 1994 and 2015.	4
Figure 1.4: Sand fly vector of <i>Leishmania</i> spp.	8
Figure 2.1: 18S, ITS1 primer composition and adapter sequence strategies work.	12
Figure 2.2: Serial dilution for Leishmania DNA.	14
Figure 3.2: PCR analyses of 18S rRNA gene of different sandflies used as in appendix 1.	27
Figure 3.3: ITS1 with 403 bp band size, and 279 bp for 18S that measured by TapeStation Analysis Software.	27
Figure 3.4: DNA sequencing for some Sand fly species as representative identification probes. Red dot: identical nucleotide, blue and black is the differences, dash is indels	28
Figure 3.6: ITS1 result of different Leishmania samples used as in appendix 3.	31

List of tables

Table 1.1: Leishmania species that cause Old world leishmaniasis.	5
Table 2.1: ITS1, 18S Amplicon PCR forward and reverse primers sequence.....	13
Table 2.3: PCR reaction mix for 18S rRNA.	15
Table 2.4: PCR indices mixture for barcoding each amplicon.	16
Table 2.5: Multiplex PCR reaction mix for ITS1 and 18S rRNA.	17
Table 2.6: ITS1 PCR mixture.	18
Table 2.7: PCR reaction mixture.	19
Table 2.8: 18S rRNA PCR reaction mixture.	20
Table 2.9 : multiplex PCR mixture.	21
Table 2.10: specific sequence virtual probes for sand flies.	23
Table 2.11: Species specific sequence virtual probes for Leishmania.....	24
Table 3.1: Next generation sequencing analysis result.	29
Table 3.2: NGS analysis result for eight Leishmania samples.	32
Table 3.5: Cross tabulation between NGS method for Phlebotomus genera and traditional method result.	37
Table 3.6: Cross tabulation between NGS method for Sergentomyia genera and traditional method result.	37
Table 3.7: Frequency of sand flies that have leishmania parasite.....	38
Table 3.8: Cross tabulation between NGS and traditional method result for Leishmania	39

List of appendices

Appendix 1: Different reference sand fly species used in the study.	50
Appendix 2: Different reference sand fly species with their gene bank accession numbers used in the study.	51
Appendix 3 : Different Reference leishmania species with WHO codes used in the study.	52
Appendix 4: NGS analysis result for 173 sand flies samples.	53
Appendix 5: work flow for NGS data analysis.....	60

List of abbreviations

WHO	World Health Organization
CL	Cutaneous leishmaniasis
MCL	Mucocutaneous leishmaniasis
VL	Visceral leishmaniasis
L	Leishmania
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
dmPCR	Direct Multiplex PCR
rRNA	Ribosomal ribonucleic acid
NTD	Neglected tropical diseases
DCL	Diffuse cutaneous leishmaniasis
PKDL	Post-kala-azar dermal leishmaniasis
LCL	Localized cutaneous leishmaniasis
Spp	Species
°C	Degree Celsius
ITS1	Internal transcribed spacer 1
SFNGSF	Sand fly next generation sequencing Forward primer
SFNGSR	Sand fly next generation sequencing Reverse primer
H	Hour
mA	Milliamps
NGS	Next generation sequencing
UV	Ultraviolet
LRU	Leishmaniasis Research Unit
TE	Tris Ethylenediaminetetraacetic acid
EB	elution buffer

Abstract

Leishmaniasis is a disease caused by *Leishmania* protozoa. There are different disease forms of leishmaniasis affecting people. The disease is transferred by the bite of infected sandflies. Identification of the *Leishmania* and sand fly species, especially in endemic areas, is important for Leishmaniasis disease control. Sand flies consist of more than 500 species, but only a few are medically important. Up until now many studies used the traditional method for sand fly species identification which relies upon morphological taxonomy of the phlebotomine Sandflies. This method has many disadvantages, for example, it requires expert entomologists to differentiate between the morphological features of sand fly species to avoid erroneous classification. Furthermore, special storage conditions for samples are needed, and the process is time consuming when dealing with large sample sizes, while delicate handlings during dissection is necessary. Until now no vaccines or safe drugs to prevent leishmaniasis infection are available, but by improving the system of sand fly and *Leishmania* parasite identification, progress in leishmaniasis disease control can be achieved.

The aim of the current study was to differentiate between the most common species of sand fly and to detect *Leishmania* DNA within the sand fly for species identification with high specificity and sensitivity using next generation sequencing. Sand flies (171) were collected from Tubas district, northern Palestine, using CDC light traps. Genomic DNA was extracted from all of them, and universal a multiplex PCR assay was setup up for sand fly 18S and Leishmanial ITS1 gene amplification. The PCR was designed to allow subsequent use of NGS Illumina platform adaptors for DNA sequencing. International reference strains of *Leishmania* and Sandflies were used in NGS system optimization for the NGS data were analyzed using galaxy online bioinformatics program, which showed the system ability to identify all reference (9) and collected sandflies (171) including the two genera: *Phlebotomus* (94.1%) and *Sergentomyia* genera (5.9%). *Phlebotomus sergenti* sand fly was the dominant species in our collection (86%).

The results were in accordance to the classical microscopic method with p value <0.001. *L. tropica* was identified in (8/171) 4.7% in the collected sandflies. Previously it was reported that Tubas is endemic region for *L. tropica* parasites.

In conclusion, the method is able to perform satisfactory high-throughput screening in ecological samples. These results will help in detecting the transmission of several

potential vectors that vary in their spatial and geographical distribution, which could explain the high prevalence of Leishmaniasis cases in specific endemic regions.

Chapter One:

1. Introduction

1.1. Leishmaniasis

Leishmaniasis is a Neglected Tropical Diseases (NTD) caused by a parasite and spread by the bite of infected sand flies. There are three major clinical forms for Leishmaniasis disease, cutaneous leishmaniasis (CL), visceral leishmaniasis (VL) and mucocutaneous leishmaniasis (MCL) (Steverding, 2017; Sundar & Rai, 2002). The most common type and least fatal form is CL, it is characterized by ulcerative skin lesions on site of insect bite, and it is caused by *Leishmania major*, *L. tropica*, *L. aethiopica*, *L. peruviana*, *L. guyanensis*, *L. panamensis*, *L. mexicana*, *L. braziliensis*, and *L. amazonensis* (Salam, Al-Shaqha, & Azzi, 2014).

The second form is visceral leishmaniasis or kala-azar in Asia, the most severe form and mostly fatal in developing countries if untreated (Desjeux, 2001). Africa, mainly Ethiopia and East Africa like Kenya, Sudan, Uganda have the highest number of VL caused by *L. infantum*. (Berman, 2006). This type caused by *L. donovani* and *L. infantum* parasites. The parasite spread to internal organs in infected patient like the liver, spleen, and bone marrow. Clinical symptoms of VL include fever, splenomegaly, hepatomegaly, progressive anemia, substantial weight loss, pancytopenia, and increase levels of a certain

immunoglobulin in the blood (Hypergammaglobulinemia) which complicated by serious infections (Sundar & Rai, 2002). The disease is fatal if not early diagnose and treated well.

Mucocutaneous Leishmaniasis (MCL) is an uncommon form of the disease and it's similar to the cutaneous form but MCL is characterized by destruction of mucous membranes of the nose and mouth, and the symptoms usually need time to appear on patient between one and five years after the skin lesions. It is caused by *L. panamensis*, and *L. braziliensis* (Salam et al., 2014).

1.2. Leishmaniasis worldwide

There are around 12 million patients suffer from leishmaniasis in the world, according to the World Health Organization – WHO, and more than 300 million people are at risk of infection (Fokialakis et al., 2007). In each year 200,000-400,000 VL cases worldwide reported and > 90% of these cases found in six countries: India, Bangladesh, Brazil, Ethiopia, Sudan, and South Sudan. The number of deaths reach to 10% after diagnosed with VL (Alvar et al., 2012). Half of these cases found in India alone (Sundar et al., 2001). The estimated annual incidence of CL is around 0.7-1.3 million new cases in worldwide and more than 95% of cases are found in the Mediterranean basin, the Middle East, Americas and Central Asia Figure (1.1). In 2015 over two-thirds of new cases found Especially in Afghanistan, Brazil, Algeria, Colombia, the Islamic Republic of Iran and Syria (Alvar et al., 2012; Organization, 2018, March 14). Multiple factors make the disease out of control including lack of effective vaccines, increasing travels between countries, difficulties in vectors control, lack of awareness, and drug increasing resistance (Khraiweh et al., 2016).

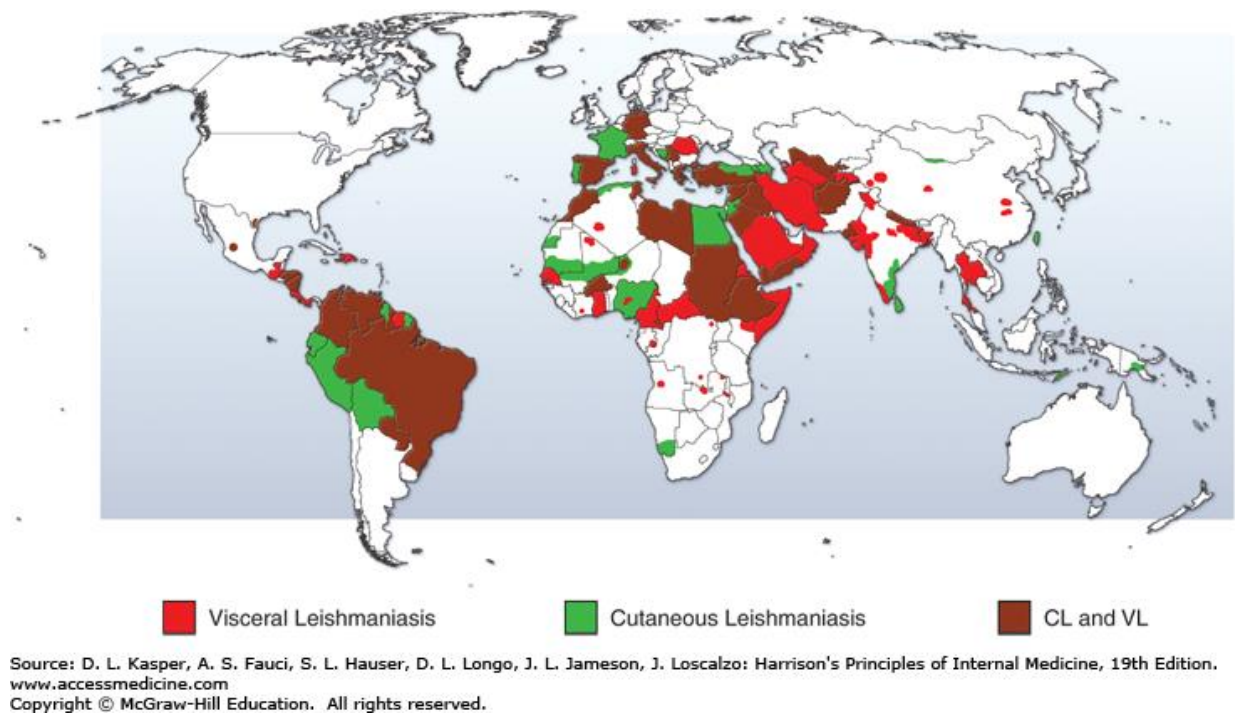


Figure 1.1: The epidemiology of leishmaniasis diseases worldwide.

1.3. Leishmaniasis in Palestine

As in other countries in the Middle East, the most common type of Leishmaniasis found in Palestine is CL, and a recent study reported 2160 CL cases clinically suspected with CL between 1994 -2015. Using molecular tests, only 895 of these cases confirmed CL with percentage 41.4%. The CL cases were collected from 77 Palestinian localities, which included cities, refugee camps, villages and Bedouin encampments. Jericho showed the highest number of CL cases with percentage more than 70%, *L. tropica* parasites being the causative agent of CL in 64 localities which cover all 11 West Bank districts, also *L. major* reported from 20 localities which cover only nine of the West Bank districts. For Gaza Strip there is no data about leishmaniasis disease (Al-Jawabreh et al., 2017; Sawalha, Shtayeh, Khanfar, Warburg, & Abdeen, 2003). Figure 1.2 shows the distribution of cutaneous leishmaniasis in the Palestinian Authority areas.

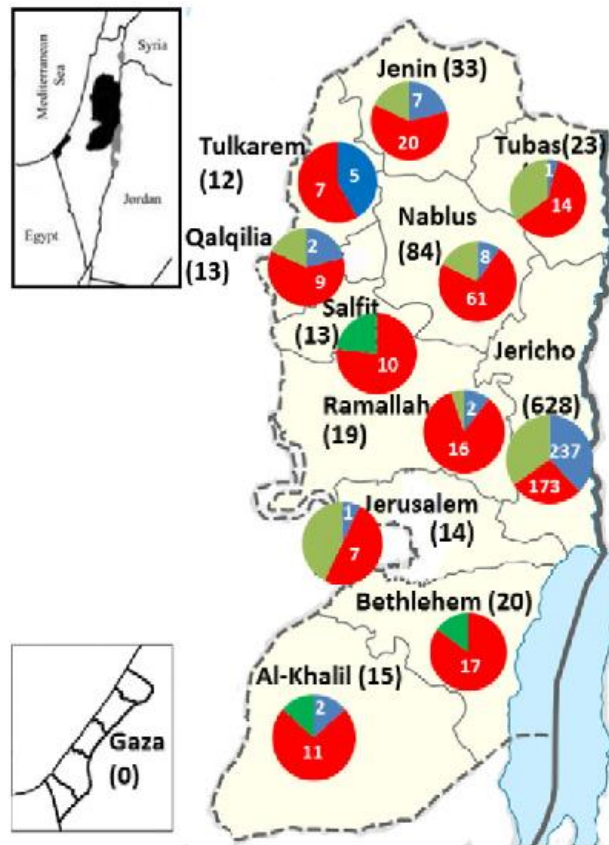


Figure 1.2: Geographical distribution of leishmaniasis disease in the Palestinian West Bank region between 1994 and 2015: numbers of cases caused by *L. major* are in blue; numbers of cases caused by *L. tropica* are in red; numbers of cases where the species of Leishmania remained undetermined are in green (Al-Jawabreh et al., 2017).

1.4. Leishmaniasis treatment.

Antimonial is the first drug line against all forms of leishmaniasis. Other different old drugs were used as: Amphotericin B, Pentamidine, and newer drugs like the Imidazoles, Paromomycin, Miltefosine, and Liposomal amphotericin B (Khraiweh et al., 2016), but these drugs have different side effect such as, toxicity, pain at injection site, renal insufficiency, long treatment time, and high cost (Rocha, Nonato, Guimarães, de Freitas, & Soares,

2013).

Recently these drugs become complicated by resistance introduction and their sensitivity in some leishmaniasis species that sometimes rely on the patient immune system status and infected leishmania species. So the treatment plan will based on risk-benefits ratios and

depending on stage of disease and this make the decision to choose best medication is not easy for physician and no drug can be used for all species of Leishmaniasis until now.

1.5. Leishmania Parasite

The genus *Leishmania* are protozoan parasites causing different forms of diseases called Leishmaniasis (Dostálová & Volf, 2012). *Leishmania* parasite is a part of Trypanosomatidae that belonging to Kinetoplastida order. Until 2017 there has been 18 different leishmania species identified as a pathogen for human, most types of parasite present in old world found in table (Table 1.1) (Steeverding, 2017).

Genus	Species	Clinical disease	Distribution
<i>Leishmania</i>	<i>L. aethiopica</i>	LCL, DCL	East Africa (Ethiopia, Kenya).
	<i>L. donovani</i>	VL, PKDL	Central Africa, South Asia, Middle East, India, China.
	<i>L. infantum</i> (syn. <i>L. chagasi</i>)	VL, CL	Mediterranean countries (North Africa and Europe), Southeast Europe, Middle East, Central Asia.
	<i>L. major</i>	CL	North and Central Africa, Middle East, Central Asia.
	<i>L. tropica</i>	LCL, VL	North and Central Africa, Middle East, Central Asia, India.

Table 1.1: *Leishmania* species that cause Old world leishmaniasis. Post-kala-azar dermal leishmaniasis (PKDL), Localized cutaneous leishmaniasis (LCL), diffuse cutaneous leishmaniasis (DCL).

1.6. Parasite Life cycle

Leishmania parasite has two morphological forms in its life cycle: the intracellular amastigote found in the mammalian host and the promastigote in the sand fly (Wheeler, Gluenz, & Gull, 2011). The life cycle starts when a parasite infected female sand fly sucks a blood meal from the human body. During that, it injects the *leishmania* parasite as promastigote form which is the infective stage enters into the skin, after that, macrophages will be activated and start phagocytosis the promastigote, then the promastigote transformed into second stage which is the amastigote, which known as the diagnostic stage. Amastigotes start multiplication and escape of macrophages to infect other cells in the body, the cycle continue until a second sand fly sucks a blood meal from infected patient and carry macrophages with amastigote and ingested of parasitized cell, after that the amastigote transform into a promastigote stage inside the midgut of sand fly then migrate to proboscis part (Figure 1.3).

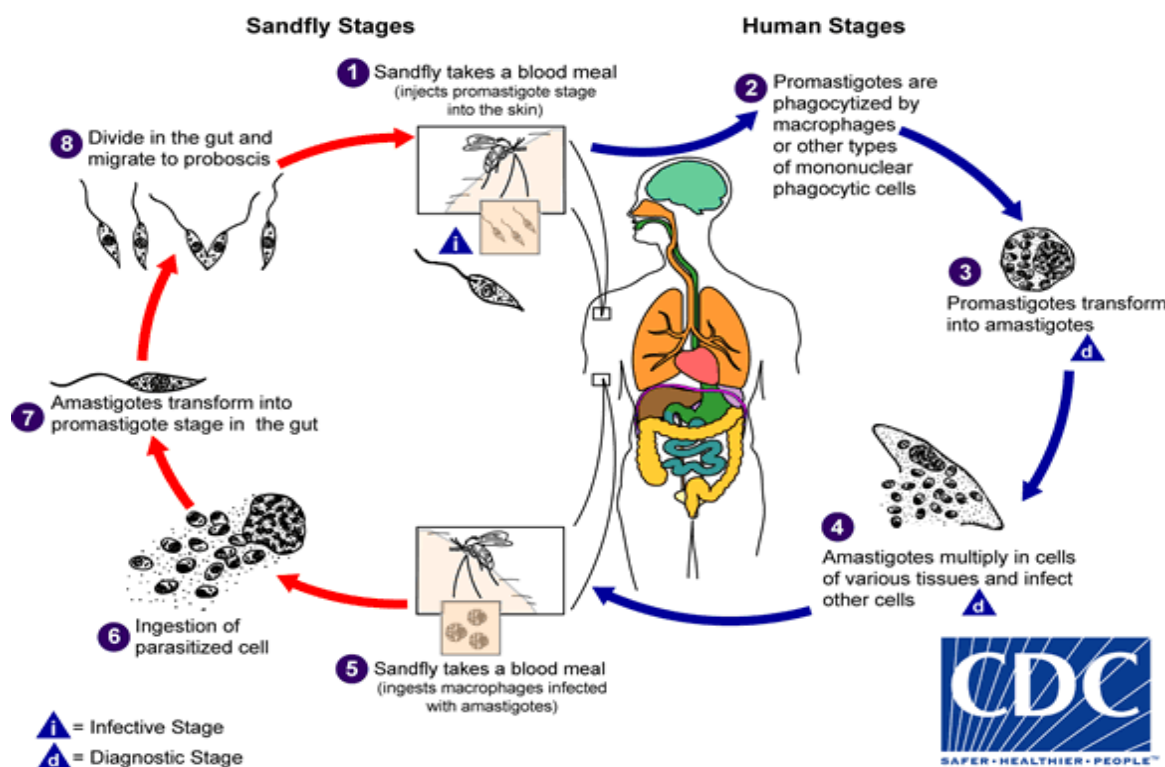


Figure 1.3: Life cycle of *Leishmania* parasite.

Leishmania can be anthroponotic (transmission from human to human or to animals) as *L. donovani* in India and *L. tropica* in the Middle East. It also can be zoonotic transmitted from wild mammal reservoir hosts to human, as rock hyraxes (*Procavia capensis*) serving as probable reservoir hosts for *L. tropica* or dogs for *L. infantum* and Sand rat for *L. major* (Alemayehu & Alemayehu, 2017). In addition, other animals were reported as reservoirs for Leishmania parasites such as: mongoose, rodents, foxes, cats, jackals, bats, wolves, primates and other domestic animals. All of these animals have been recorded as hosts or reservoirs of Leishmania species, and can maintain the transmission of parasite in different regions of the world (Alemayehu & Alemayehu, 2017).

Recent regional study suspect that rock hyraxes (*Procavia capensis*), is a reservoir animal for *Leishmania tropica*, and they found that hyrax have antibodies against *L. tropica* (Talmi-Frank et al., 2010). In addition, scientists managed to isolate one strain of *L. tropica* from infected hyrax northern Palestine (Svobodova et al., 2006).

1.7. Sand Fly

Sand fly, insect belongs to the Phlebotomidae family under the Diptera order (Bates, 2008). In Theodor's classification system Sand fly have been classified into two groups: Old world contains two genera-*Phlebotomus* and *Sergentomyia*, and the New World contains *Lutzomyia* genera (D. Lewis, 1971). Sand fly adults are small flies have a length about 3 mm, golden, brownish or gray colored. Female sand flies are blood feeders needed to produce eggs, unlike the males that does not feed on blood and feed on plant juices and sugary secretions (Figure 1.4).



Figure 1.4: Sand fly vector of *Leishmania* spp.

Female phlebotomine sand flies are the vectors of leishmaniasis and responsible for affecting millions of people in more than 80 countries. Sand fly also considered as vectors for other diseases as bartonellosis, orbiviruses and flaviviruses that cause different health problems for animals and humans (Alexander & Maroli, 2003). Interesting selective relationship between parasite species and sand fly species was observed, so specific *Leishmania* species of medical importance are usually transmitted by only one or two particular sand fly species, for example: *Phlebotomus perfliewi* is a vector for *L. infantum* that cause visceral leishmaniasis, and *Ph. papatasi* transmitting *L. major* the causative agent of CL (Killick-Kendrick, 1999), also *Ph. sergenti* is vector for *L. tropica* (Al-Jawabreh et al., 2017). So finding these specific vectors will contribute strongly in disease elimination through applying prevention control programs.

Leishmania species and vector female sand fly identifications with different techniques are highly important for disease management and the planning of control strategies within endemic areas. Morphological identification of Sand fly species can be performed by standard entomological aspects based on individual sand-fly anatomy of the head and the genitals under microscope, expert entomologists is needed for identification of different features like spermatheca, cibarium, genitalia, and the antennae of the sandflies (Giantsis, Chaskopoulou, & Claude Bon, 2017).

Use of Morphological method has many limitations such as time consuming, especially when dealing with high number of sand fly samples, need specific technical skills in sample preparation and identification and this lead to epidemiological errors.

Further and newer techniques can be used like biochemical techniques such as enzyme electrophoresis, gas chromatography and DNA probes (Kassem, Fryauff, El Sawaf, Shehata, & Shoumar, 1990; Moore et al., 1987; Ready, Smith, & Killick-Kendrick, 1988). But these techniques have limitations like, expensive cost, need longer time, less specific.

1.8. Molecular technology

Currently, with the advances in molecular biology, several genes are used to distinguish sand fly or leishmania parasite. Ribosomal RNA (*rRNA*) *18S* nuclear gene and mitochondrial DNA (mtDNA) cytochrome b (*cytb*) gene which used for sand fly identification (Bounamous, Lehrter, Hadj-Henni, Delecolle, & Depaquit, 2014). For leishmania parasite identification, ITS1 region of ribosomal RNA gene is a specific target gene (Schönian et al., 2003). There are several techniques used to identify sand fly or leishmania like polymerase chain reaction (PCR) which a simple and valid molecular tool. Other techniques like Direct Multiplex PCR (dmPCR), and polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) (Giantsis et al., 2017).

Until recently, Sanger sequencing method was the most widely used sequencing method, that have limitations like, time-consuming, expensive when done in large scale, and need a large amount of template DNA for each read. Therefore, the demand for faster, more accurate, and more cost-effective method has led to the development of NGS methods. NGS methods are high-throughput technologies with capabilities of sequencing large numbers of different DNA sequences at once (Basho & Eterovic, 2015).

1. 9. Study Objectives

1. To develop specific and sensitive High-throughput screening system using New Generation sequencing (NGS) for sand flies and leishmania parasite identification.
2. To compare the introduced method to the traditional microscopic methods.
3. To determine Leishmania parasite infection within sand fly species.
4. To find the sand fly regional abundance and their infection rate with Leishmania parasite in Northern Palestine.

Chapter Two:

2. Materials and Methods

2.1. Sand fly sample collection

Most of phlebotomine sand flies samples in this study were collected from Tubas district in North of Palestine by using CDC light traps. Isolated sand flies from traps were washed with 2% detergent solution and then stored in 70 % alcohol. Morphology and taxonomic keys of sand fly based on different taxonomic keys were used (Kakarsulemankhel, 2010; Lane, 1986; D. J. Lewis, 1982), these keys help to identify and separate blood-fed female sand flies with a help from (entomologist Mr. Samer Sawalhah, Palestinian entomology department Ministry of health) (Sawalha, Ramlawi, Sansur, Salem, & Amr, 2017). After dissection, we separated the abdomen and the thorax from sand flies, then immersed samples in ethanol and individually stored in a freezer (-20 °C).

2.2 DNA Extraction

Genomic DNA was extracted from the abdomen and the thorax of sand flies specimens manually by phenol method, this method has been the dominant and effective method of isolating genetic material for several years and give better performance than commercial kit extraction protocols (Casaril et al., 2017). We start by adding 200 uL of DNA lysis

buffer (50 mM NaCl, 10 mM EDTA, 50 mM Tris-HCl pH 7.4, 1% Triton X-100) to sand fly which were stored in alcohol, then we put the sample on vortex and use a wooden stick to make a homogenized sample. Proteinase K (20 uL, 10 mg/mL) was added and incubated at 60 °C using a mixer (VorTemp™ 1550) for one hour until complete digestion of tissue. Phenol solution (PH:8) (0.2 ml) was added to samples and vortexed for 1 min, then centrifuged for 3 minutes at 14000 rpm, and the aqueous layer which contains the DNA was collected in new 1.5 ml Eppendorf tube. Sodium acetate (8 µL, final conc 0.2 M) and 700 µL of cold 100% ethanol were added. Then samples were incubated at -70°C for 30 minutes to precipitate DNA. The samples were then centrifuged for 10 minutes at 14000 rpm, at 4°C. The supernatant was discarded, and the pellet was washed with 0.3 ml of cold 70% ethanol, followed by centrifugation for 5 minutes at 14000 rpm, at 4°C. Following ethanol removed with temperature using speed vacuum centrifuge (Concentrator Plus™), the samples were re-suspended in 50 µL of 1× TE (Tris Ethylenediaminetetraacetic acid). DNA concentration was measured by Nanodrop 2000 and assessment of nucleic acid purity, an interpretation of 260/280 ratios and 260/230 ratios was measured, samples then kept frozen at -20 °C until further use.

2.3 Design of NGS specific primers

There are many target genes used in PCR for Leishmania detection and identification, as examples: kinetoplast DNA (kDNA), Internal transcribed spacer 1 (ITS1), the gp63 and the miniexon (spliced leader RNA) gene (Monroy-Ostria, Nasereddin, Monteon, Guzmán-Bracho, & Jaffe, 2014). ITS1 region of ribosomal RNA gene was selected based on its high sensitivity, specificity and its ability to identify Leishmania species, furthermore, hundreds of studies used this gene to detect leishmania parasite and determine the species within different clinical sample types (Dávila & Momen, 2000; Schönian et al., 2003).

The specific forward and reverse primers for amplicon PCR ITS1- rRNA gene were designed to amplify 276 bp segment of the conserved region of the ITS1 gene of Leishmania. In addition, Illumina adapter overhang nucleotide sequences are added to the gene- specific (ITS1) sequences as in table (2.1). These adapter sequences must be appended to the primer pair sequences for compatibility with Illumina index and

sequencing adapters which essential for Illumina platform sequencing systems as in figure (2.1).

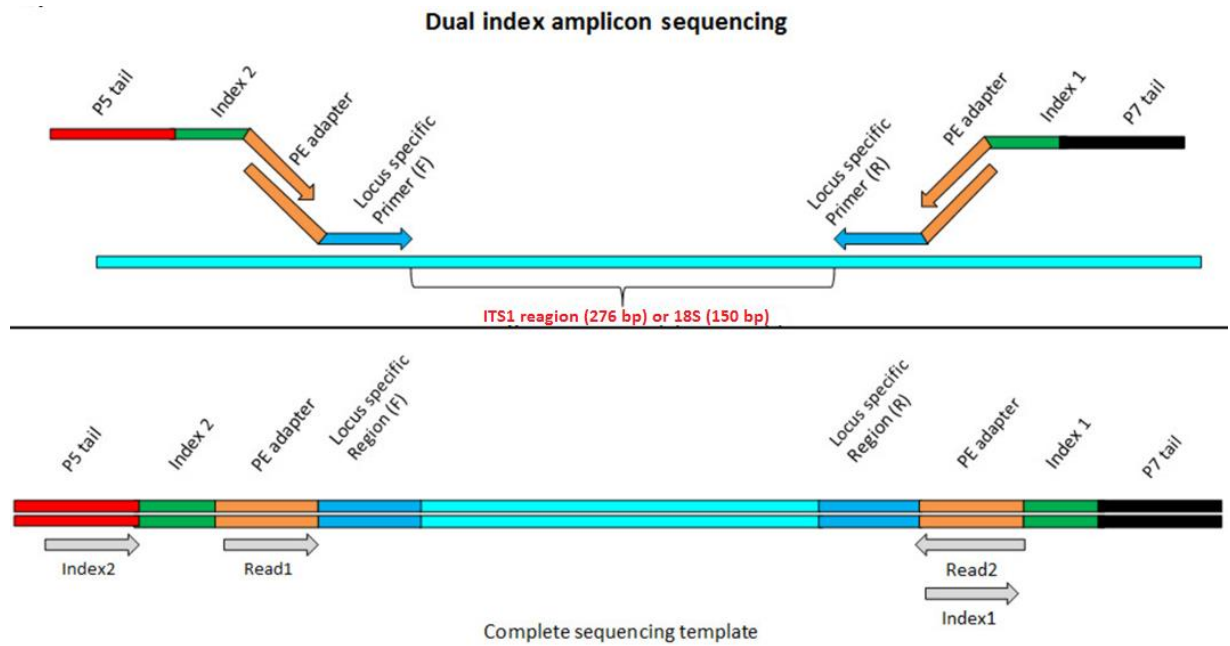


Figure 2.1:18S, ITS1 primer composition and adapter sequence strategies work.

For sand fly PCR, universal primers for all sand fly species including *Phlebotomus* and *Sergentomyia* rRNA gene were included, reverse universal primer for Amplicon 18S rRNA gene was designed to amplify 150 bp (fit with Illumina sequencing kit), the forward Amplicon primer for 18S was used from a previous study (Giantsis et al., 2017). Overhang Illumina adapter sequences were added as previously described.

Table 2.1: ITS1, 18S Amplicon PCR forward and reverse primers sequence.

	Illumine adapter sequence	Primer sequence
ITS1Amplicon PCR Forward Primer	(5'- TCGTCGGCAGCGTCAGAT GTGTATAAGAG -3')	(5'- ACAGAGCTGGATCATTTTCCGA TG -3')
ITS1Amplicon PCR Reverse Primer	(5'- GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAG -3')	(5'-ATCGCGACACGTTATGTGAG AACGGCTCACATAACGTGTCGCGA -3')
18S rRNA Amplicon PCR Forward primer	(5'- TCGTCGGCAGCGTCAGATG TGTATAAGAGACAG -3').	(5'- TGCGGTTAAAACGTTCGTAG -3')
18S rRNA Amplicon PCR Reverse primer	(5'- GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAG -3').	(5'- ACCGGTAAAACATCCGTCAC -3').

2.4 Amplicon PCR Amplification

Conventional PCR was performed on all collected samples using PCR tubes with a 25 μ L reaction mixture containing 5 μ L of DNA template and 20 μ L of Distilled H₂O, Forward and Reverse primers (0.8 μ M each one), X2 PrimeSTAR Max Ready Mix (TAKARA). PCR tube with only water and the previous mix was included as negative control.

PCR Amplification was performed on T100™ Thermocycler (BIORAD) using the following amplification protocol: initial denaturation for 5 minutes at 98°C, Thermo cycle file with 35 cycles of denaturation for 10 seconds at 98°C, annealing for 5 seconds at 55°C, and extension for 10 seconds at 72°C, then final extension for 5 minutes at 72°C.

2.5.1 Leishmania primers optimization

International reference strain of *Leishmania tropica* LRC-L590 (WHO strain identifier MHOM/IL/1990/P283) was used as DNA sample a positive control. DNA concentration was 46 ng/ μ L, and serial dilutions was done for DNA with Ultra pure water molecular grade (UPW) and dilution was started with dilution factor 1:5 by adding 2 μ L from stock DNA to 8 μ L in the first tube then 2 μ L from Tube 1 was transferred to 8 μ L of diluent in

Tube 2 and mix. Then transfer 2 μL from Tube 2 to next Tube and mix. This process was repeated until we have eight serial dilutions tubes figure (2.2).

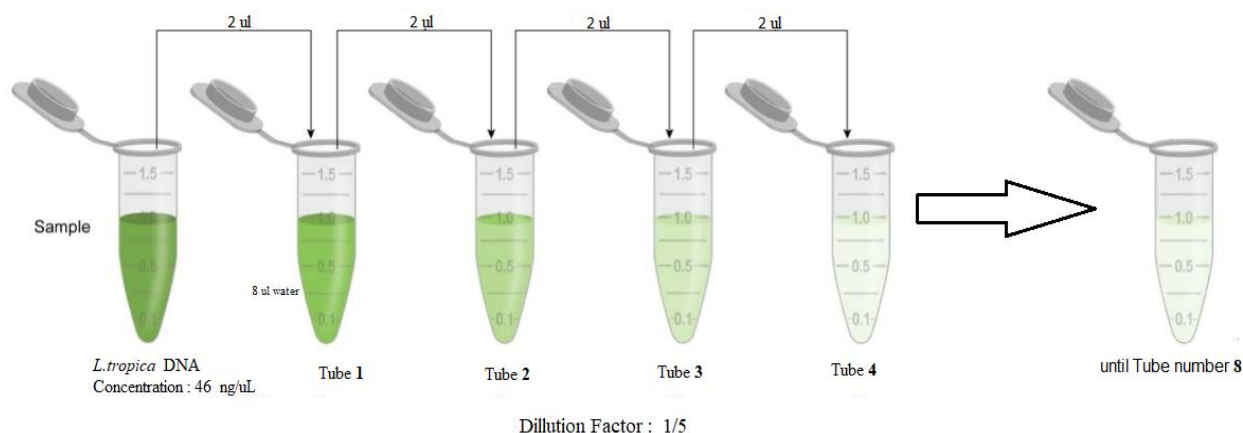


Figure 2.2: Serial dilution for Leishmania DNA.

PCR reaction mixture was prepared for nine tubes as in table (2.2), and added the diluted DNA to PCR mixture, in the last tube, also 2 μL UPW was added as a negative control.

Table 2.2: PCR reaction mixture for ITS1.

		Volume (μL)
1	DNA	2
2	ITS1Amplicon PCR Forward Primer 100 μM	0.2
3	ITS1Amplicon PCR Reverse Primer 100 μM	0.2
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
5	UPW	10.1
Total volume		25

Then PCR in T100™ Thermocycler was performed using the program which explained in previous. After that, we take 7 μL from each amplified product and mix it with 1.5 μL

Thermo Scientific 6X DNA Loading Dye and loaded in 2% agarose gel and compared with a 100-bp DNA Ladder (Invitrogen). Electrophoresis gels were run in 1X TAE buffer at 120 V, 80 mA, 10 W for 1 h and stained with ethidium bromide and visualized under UV transillumination by using DNR Bio Imaging Systems.

2.5.2 Sand fly primer Optimization

Optimized sand fly primers started with eight DNA samples from different characterized sand fly species, according to Appendix (1). We measured the DNA concentration and make a dilution of samples that have high concentration to reach until (10ng/ul). PCR reaction mix was prepared as in table (2.3), and added 5 μ L of DNA to each PCR reaction also added UPW as a negative control in the last tube.

Table 2.3: PCR reaction mix for 18S rRNA.

		Volume (μ L)
1	Sand fly DNA	5
2	18S rRNA Amplicon PCR Forward Primer 10 uM	1.0
3	18S rRNA Amplicon PCR Reverse Primer 10 uM	1.0
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
5	UPW	5.5
Total volume		25

PCR was performed in T100™ Thermocycler using the program and gel documentation was done as above.

2.5.3 PCR products purification and NGS

After PCR for Leishmania and sand flies were ready, both products were pooled, for example, PCR product 1 of leishmania sample number 1 was mixed with PCR product from sand fly number 1, respectively. After that mixed PCR was purified using AMPure

XP beads (GE Health Backman Coulter, Jerusalem, Palestine) according to a company manual sheet, by adding 25µL of UPW to 25µL of PCR product, so starting was with total volume 50 µL, then 50 µL of magnetic beads were added to each well and incubate for 5 min at RT. Wells were transferred to a magnetic field for 3 min , then 95 µL of supernatant was removed slowly by pipette without touch the pellet. After that, we added 200 µL of freshly prepared 80% ethanol (4 ml absolute ethanol + 1 ml UPW) and removed after 30 sec and remain the wells cover open for few minutes to allow ethanol evaporation. Finally, we added 30 µL of EB and leave wells out of the magnetic field for 2 min, then return it to magnetic fields for 5 min and pure DNA was transferred into new fresh tube.

To discriminate between samples we added dual- index barcodes by a second PCR as in table (2.4). These barcodes partially compatible with first added Illumina primers that were over hanged with the target genes of Leishmania and sandflies.

From each PCR product, 15 ul was collected from 8 tubes in one pool tube and secondary PCR purification was applied using AMPure XP beads. The product was eluted in 120 ul and is ready to run in Illumina next generation sequencing template.

Table 2.4: PCR indices mixture for barcoding each amplicon.

		Volume (µL)
1	Purified PCR1 product	7.5
2	Nextera XT Index Primer 1 (N7xx) 5 uM	2.5
3	Nextera XT Index Primer 2 (S5xx) 5uM	2.5
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
Total volume		25

2.6 Multiplex PCR and Sand fly primers as limit factor

The purpose of multiplex PCR is to do the ITS1 and 18S PCR in one PCR reaction. This will save time, DNA and costs. The main problem is the extra Sand-fly DNA in a sample and this could consume the PCR contents and lower the Leishmnaia ITS1 PCR sensitivity,

to overcome this problem, limiting factor for Sandflies 18S PCR amplification must be controlled.

To set up sand fly PCR primers as a limit factor, multiplex PCR for both leishmania and sand fly was set up using the same previous conditions, with different serial dilution of sandflies primers for each reaction. This will limit the minimal amount of primers requested for sandflies PCR reaction. Mix of 1 μ L of SFNGSF (10uM) and 1 μ L of SFNGSR (10uM) to 8 μ L of UPW and mix, after that 2 μ L was transferred to the second tube that has 8 μ L UPW and continues as a serial dilution to tube number 6.

Also, we prepared the mix for multiplex PCR as in table (2.5) for 8 tubes and added 1 μ L of SFNGSF and 1 μ L of SFNGSR with 10 uM to tube number 1 and continue adding 2 μ L from sand fly primer serial dilution to each tube until tube number 7 and last tube number 8 does not have sand fly primer as negative control.

Table 2.5: Multiplex PCR reaction mix for ITS1 and 18S rRNA.

		Volume (μL)
1	Sand fly DNA	0.2
2	Leishmania DNA	0.2
3	ITS1Amplicon PCR Forward Primer 100 uM	0.25
4	ITS1Amplicon PCR Reverse Primer 100 uM	0.25
5	SFNGSR 10 uM	1.0
6	SFNGSF1 0 uM	1.0
7	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
8	UPW	9.6
Total volume		25

Then PCR was performed in a thermal cycler using the same program previously described above. Then 7 μL from each amplified products was loaded on the gel and captured as previously described above.

2.7 Selection of Leishmania species specific probes

Optimized Leishmania primers by using nine of DNA samples for different worldwide leishmania species that were identified previously on the species level, see Appendix (3), we measured the DNA concentration and performed PCR as in table (2.6).

Table 2.6: ITS1 PCR mixture.

		Volume (μL)
1	Leishmanial DNA	2
2	ITS1Amplicon PCR Forward Primer 100 uM	0.2
3	ITS1Amplicon PCR Reverse Primer 100 uM	0.2
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
5	UPW	10.1
Total volume		25

Then perform PCR in T100™ Thermocycler using the program which was described previously. Then 7 μL from each amplified products was loaded and visualized under UV transillumination using DNR Bio Imaging Systems as previously described.

PCR mix purification was done by using AMPure XP beads as previously described. Then second barcoded indecies PCR and purified as previously described. Final pooled library quantification was done using Qubit machine (Invitrogen) and normalized to 4 nM, then send for next generation sequence (NGS) targeting >20000 reads per sample.

2.8 Optimizing the quality copy number for Leishmania

We use the previous sample for *L. tropica* LRC-L590 strain, their DNA concentration was 46 ng/uL and we make serial dilutions for DNA with UPW in eight tubes with Dilution factor (1:10).

PCR reaction mixture was prepared for eight tubes as in table (2.7), and added the DNA from serial dilution to PCR mixture and in the last tube, 2 μ L UPW was added as a negative control.

Table 2.7: PCR reaction mixture.

		Volume (μ L)
1	<i>Leishmania</i> DNA	2
2	ITS1Amplicon PCR Forward Primer 100 uM	0.2
3	ITS1Amplicon PCR Reverse Primer 100 uM	0.2
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
5	UPW	10.1
Total volume		25

Then PCR performed in T100™ Thermocycler using the previously described program and loaded on the gel and visualized under UV transillumination by using DNR Bio Imaging Systems as described previously.

For sand fly, we use DNA sample for *Ph. serganti* their concentration was 10.5 ng/uL, we dilute DNA with ultrapure water in two tubes first one have dilution factor 1/10 and the second tube have a dilution factor 1/100. Then we prepare PCR reaction mixture for three tubes as in table (2.8), and added the DNA from serial dilution to PCR mixture and in the last tube we added 5 μ L UPW as negative control.

Table 2.8: 18S rRNA PCR reaction mixture.

		Volume (μL)
1	Sand fly DNA	5
2	18S rRNA Amplicon PCR Forward Primer 10 uM	1.0
3	18S rRNA Amplicon PCR Reverse Primer 10 uM	1.0
4	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
5	UPW	5.5
Total volume		25

PCR was done, loaded and visualized as previously described. After doing the previous PCR for Leishmania and sand flies, we continue the optimization by mix the PCR product for sand fly in one tube, we mix the PCR product for SF sample number 1 (1/10) with PCR product for sand fly number 2 (1/100), and take 3 μL from mix tube and add it to each leishmania dilution PCR product respectively for seven tubes. After that, we make DNA purification by using magnetic beads which made by GE Health Backman Coulter as in their manual sheet. Then make second index PCR before uses magnetic beads again to clean up the final library before quantification and run the samples on Next generation sequencing then analysis the result.

2.9 Analysis of field Sand flies:

Followed PCR optimization for the Leishmania and sandflies primers on DNA isolated from a reference sample, the collection of sand flies from Tubas district in Northern of Palestine was done in different three types of collection with the total number 171 sand flies, and we followed the steps as described previously. The first collection sandflies that were sorted and identified manually by microscopic and ITS1 PCR analysis. Second collection, sorted Sandflies and identified microscopic unblind. Third collection were sandflies collected directly and blindly identified by NGS. Work on sand flies samples collected from the field was applied. After doing the DNA extraction and measure the

DNA concentration for 171 samples we optimize samples with high concentration to make their DNA concentration equal 20 ng/uL and make dilution for sand fly primer by mix 100 μ L of (18S rRNA Amplicon PCR Forward Primer 10 uM) with 100 μ L of (18S rRNA Amplicon PCR Reverse Primer uM) and dissolved in 800 μ L of ultrapure water. To reach final concentration 2 uM for each sand fly primer. Then run multiplex PCR as in table (2.9), and use positive control (leishmania DNA) and negative control (ultrapure water).

Table 2.9 : Multiplex PCR mixture.

		Volume (μL)
1	Sand fly DNA	2.5
2	Sand fly Primers mix (0.8uM)	5
3	ITS1Amplicon PCR Forward Primer 100 uM	0.25
4	ITS1Amplicon PCR Reverse Primer 100 uM	0.25
5	ReadyMix (PrimeStar Max DNA Polymerase)	12.5
6	UPW	4.5
Total volume		25

PCR was done and visualized as previous described. Followed with beads purification as explained previously and eluted in 100 uL of elution buffer (Library).

Library purity and quantity was evaluated again by TapeStation machine (TapeStation Analysis Software A.02.01 © Agilent Technologies, Inc. 2015) and Qubit machine. The concentration of 4 nM was prepared from pooled samples. Twenty thousand reads for each sample were targeted. Samples were deep sequenced on NextSeq 500/550 machine using the 150-cycle Mid Output Kit (Illumina).

2.10 Bioinformatics analysis

Fastaq NGS files data that was obtained from the sequencing machines was uploaded online on data analysis platform Galaxy (<http://usegalaxy.org>) which is widely deployed and developed for next generation sequence (NGS) data analysis. Galaxy integrates many bioinformatics tools within one windows interface. Also, special workflow can be created with customized specific probes (Blankenberg et al., 2010). The workflow can deal with hundreds of samples with millions of sequences reads.

Workflow design:

The created workflow was consist of: Trim galore which a wrapper script to automate quality and adapter trimming as well as quality control, then remove sequencing artifact from data as polynucleotides as an example (GGGGGGGGGGGGGGGGGGGG). Then filter FASTQ reads by determining a minimum and maximum read length between (100-151 bp) and select a minimum quality score >20, which represents an error rate of 1 in 100 (according to Illumina Next-Seq machine sequencing error rate), with a corresponding call accuracy of 99%. That is mean specify minimum per-base quality scores, with optionally specifying the number of bases that are allowed to deviate from this range.

Then convert FASTQ sequencing reads to FASTA sequences followed with TAB-delimited format, and then to Synax tool which searches the data for lines containing or not containing specific virtual probes that already added them to a workflow. These virtual probes contain specific sequence for each sand flies species, in addition a common universal sequence which should be found in all sand fly species sequences Table (2.10). Then converts tab delimited data into FASTA formatted sequences and finally, tool is to collapse identical sequences into a FASTA file into a single sequence. The final step include reads count of the filtered sequences.

As application, the fastaq files obtained from the machine were uploading on usegalaxy.org as Fastqsanger.gz, then run the workflow as mentioned above, see appendix (5). The workflow was applied for leishmania and sandflies retrieved sequences with exception of selected probes by adding specific sequence for each species and each genotype, and common sequence for all Leishmania species with forward primer to increase the specify of the analysis Table (2.11).

Sand fly virtual Probes

Table 2.10: specific sequence virtual probes for sand flies.

	SF species	Specific virtual probe
1	Sand Fly (common)	(TGCGGTAAAACGTTTCGTAG)
2	<i>S. dentate</i>	(ACACGGGCAATGCAC)
3	<i>S. barraudi</i>	(TTTGTGCATTGGTAAAACAGTGTGCAA)
4	<i>S. minuta</i>	(GTGTGCAAATGACTTTA)
5	<i>Ph. sergenti</i>	(GCTCTGTGCGTTTTGTGTA)
6	<i>Ph. sergenti</i> G1	(GTGTAAAAACAAACGTATA)
7	<i>Ph. perfiliewi</i>	(CGCATATGTTTCACCGTA)
8	<i>Ph. perfiliewi</i>	(TACTATATGTTTCACCGTCA)
9	<i>Ph. neglectus</i>	(TCGCATATGTGTGTCTCACCGTCA)
10	<i>Ph. sergentomaya</i>	(AAAACAGTGTGCAGGTGA)
11	<i>Ph. papatasi</i>	(CTGTGCGTTCTGTGTAAAAGCAAGCGTATAGT)
12	<i>Ph. syriacus</i>	(TCGCATATGTGTGTCTCACCGTAA)
13	<i>Lutzomyia</i>	(ATTACTTGTTGCCTATG)
14	<i>Lutzomyia umbratilis</i>	(AAACAGTGTGTAT)
15	<i>Ph. argentipes</i>	(CACTGTCAAAGGTGGCGTA)
16	<i>Ph. tobbi</i>	(AAAGTGTCATATGTATGT)

Leishmania virtual Probes

Table 2.11: Species specific sequence virtual probes for Leishmania.

	Leishmania species	Specific virtual probe
1	Leishmania (common)	Forward (GATCATTTTCCGATG)
2	<i>L. major</i>	(TTTATACTCAAAATTTGCA)
3	<i>L. donovani</i>	(ATTACACCAAAAAA)
4	<i>L. tropica</i>	(CATATACAAAACTCGGGGAGGCCTAT)
5	<i>L. infantum</i>	(ATCGACGTTATAACGCA)
6	<i>L. atropica</i>	(TCGGGCAGGCCTATTA)
7	<i>L. tropica</i> G1	(ATTACACCCCAAAAAAACA)
8	<i>L. tropica</i> G2	(ATTACACCCCAAAAAAACA)

2.11. Statistical analysis and multiple sequence alignment

Statistical analysis of the results was performed using Statistical Package for the Social Sciences (SPSS) version 20.0 (IBM, Corporation) and Microsoft Excel version 2016, showing the results that presented as frequencies and p-values were analyzed by using Pearson Chi-Square. Fisher's test analysis was done using online graphpad : <https://www.graphpad.com/quickcalcs/contingency2/>. Multiple sequence alignment was done online (<http://multalin.toulouse.inra.fr/multalin/>) as described by Corpet (Corpet, 1988).

Chapter Three:

3. Results

3.1 PCRs optimization

ITS1 PCR for *Leishmania tropica* DNA dilution with a concentration of 46 ng/uL and the serial dilution was done by dilution factor (1:5), give positive clear band for first four dilutions, and product is approximately the expected size product 343 bp (ITS1 for *L. tropica* 276 bp and Illumina primers 67 bp), also the dilution sensitivity reached up to 3.2 ng/uL as shown in figure (3.1).

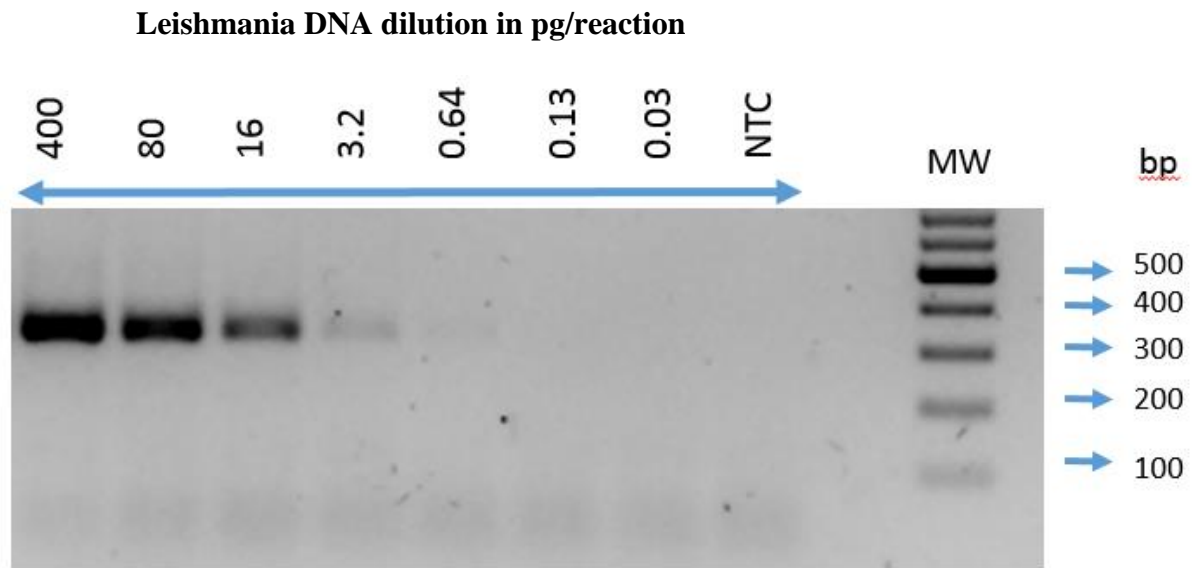


Figure 3.1: PCR analyses of ITS1 gene of DNA *L. tropica* dilutions. NTC: negative control. MW: 100 bp DNA molecular weight ladder.

For sand fly samples with high DNA concentration for eight sandflies that were identified on the species level by an entomologist, were diluted to be between 10-100 ng/ul.

PCR for the eight samples was positive and the negative control (UPW added) showed no band as shown by results of PCR product on gel electrophoresis. Amplification based on 18S rRNA primer gives one clear band with an expected size equal 217 bp figure (3.2).

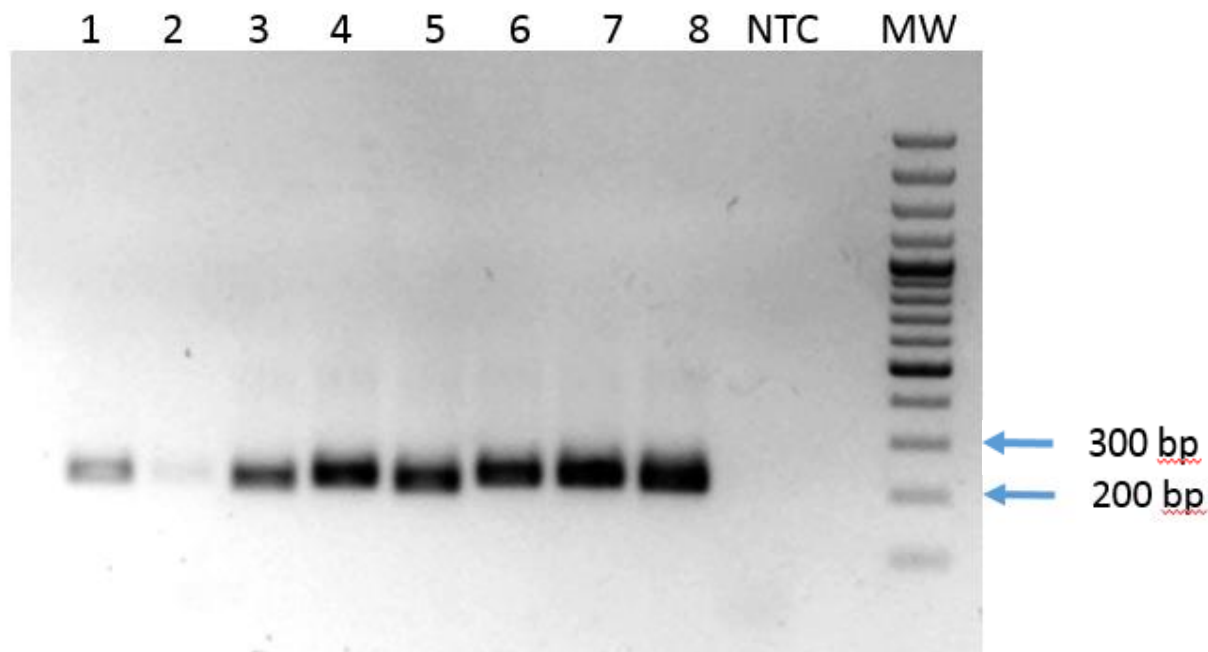


Figure 3.2: PCR analyses of 18S rRNA gene of different sandflies used as in appendix 1, **1:** *Ph. tobbi*, **2:** *Ph. argentipes*, **3:** *Lutzomyia* spp, **4:** *Ph. sergenti*, **5:** *Sergentomyia* genus **6:** *Ph. syriacus*, **7:** *Ph. perfelewi*, **8:** *Ph. papatasi*, **NTC**; negative control.

TapeStation Analysis Software A.02.01 (Agilent Technologies, Inc. 2015) showed a clear band at the 279 level (see figure 3.3), this band mainly represent the 18S amplicon for all the sandflies samples together. The ITS1 band is about 350 bp.

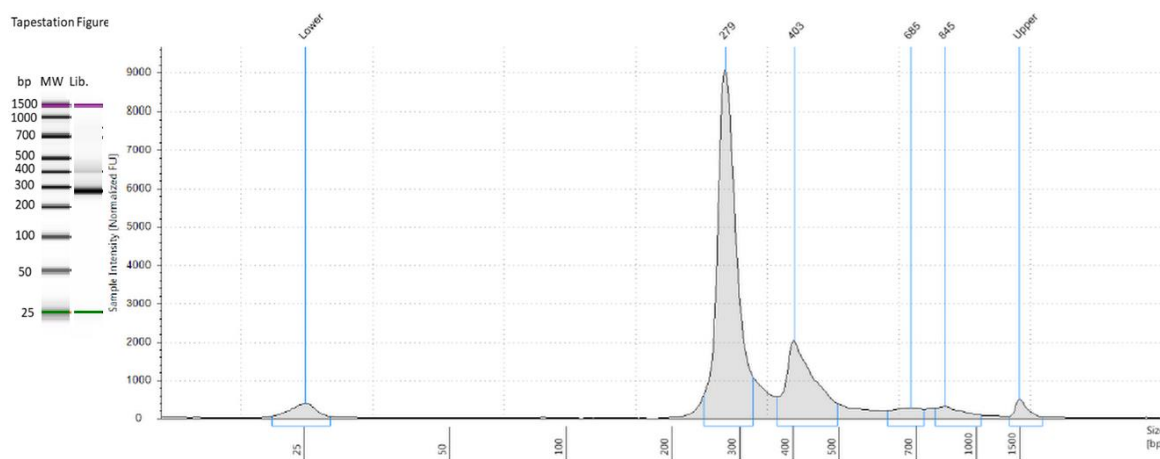


Figure 3.3: ITS1 with 403 bp band size, and 279 bp for 18S that measured by TapeStation Analysis Software.

The eight samples gave sequencing reads between 42708-134399 after the high stringency filtration using the galaxy workflow.

Multiple alignment for reference strain sand flies species sequence with obtained sand flies sample to determine specific SNP that differentiate each species shown in the following figure (3.4). The obtained sequences showed a high number of reads and they as followed: **A:** *Ph. argentipes* with reads count 45481, **B:** *Ph. syriacus* with 76229 reads, **C:** *Ph. perfelewi* with 134399 reads, **D:** *Ph. tobbi* with 42708 reads, **E:** *Ph. papatasi* with 110988 reads, **F:** *Ph. sergenti* with 82952 reads, **G:** *Sergentomyia* genus with 109246 reads, **H:** *Lutzomyia* spp with 102931 reads.

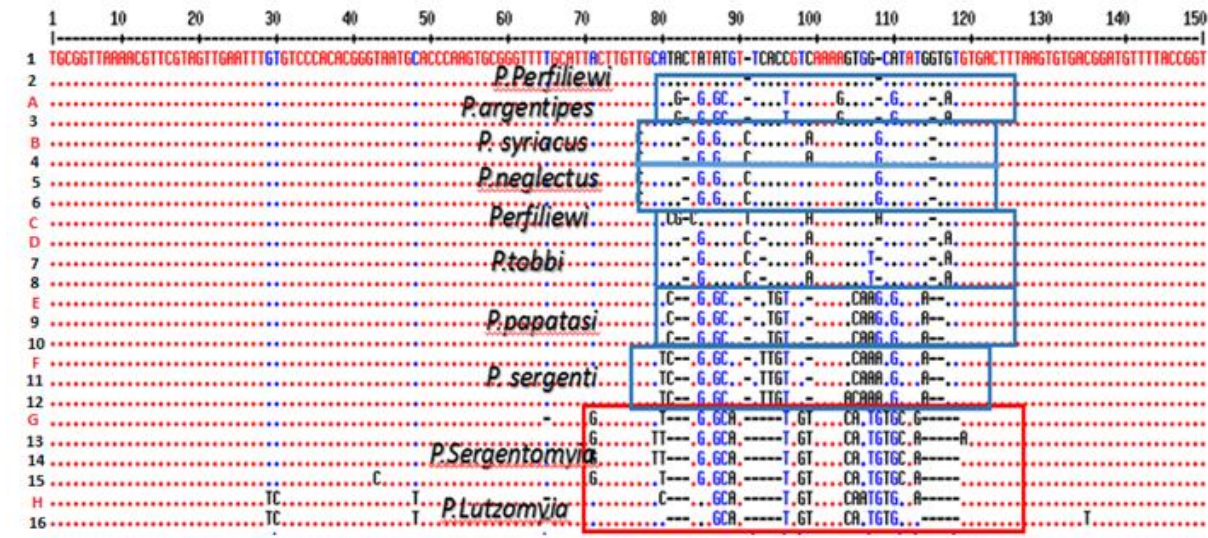


Figure 3.4: DNA sequencing for some Sand fly species as representative identification probes. Red dot: identical nucleotide, blue and black is the differences, dash is indels. Numbers from 1-16 is sample numbers for reference sand flies used See appendix 2. **1, 2:** *Ph. perfilewi* **3:** *Ph. argentipes* **4:** *Ph. syriacus* **5, 6:** *Ph. major* **7, 8:** *Ph. tobbi* **9, 10:** *Ph. papatasi* **11, 12:** *Ph. sergenti* **13:** *S. minuta* **14:** *S. barraudi* **15:** *S. buxtoni* **16:** *P. Lutzomyia*.

A-H reference samples as in appendix 1, **A:** *Ph. argentipes* **B:** *Ph. syriacus* **C:** *Ph. perfelewi* **D:** *Ph. tobbi* **E:** *Ph. papatasi* **F:** *Ph. sergenti* **G:** *Sergentomyia* genus **H:** *Lutzomyia* spp. Which obtained from entomology laboratory after sequencing on Illumina plate form.

NGS analysis was summarized in table (3.1).

Table 3.1: Next generation sequencing analysis result.

Mix number	Sand fly species	SF copy number	<i>L. tropica</i> copy number
1	<i>Ph. tobbi</i>	42708	1747
2	<i>Ph .argentipes</i>	45481	976
3	<i>Lutzomyia spp.</i>	102931	594
4	<i>Ph. sergenti</i>	82952	340
5	<i>Sergentomyia spp.</i>	109246	142
6	<i>Ph. syriacus</i>	76229	37
7	<i>Ph. perfiliewi</i>	134399	9
8	<i>Ph. papatasi</i>	110988	2

As we have shown in the result the serial dilution for Leishmania DNA present with a decrease in the read number for diluted tubes and the last three tubes can be considered as negative samples. For sand flies the reads are high but the instability of reads because their differences in the primary DNA concentration for Leishmania, and sand fly which have the dominant DNA in the mix. So to get more stable result we make multiplex PCR for Leishmania and sand fly in eight serial dilutions. All dilutions have the same content but the sand fly primer concentration was different with a dilution factor of (1:5) and negative control does not have SF primers. Clear bands for ITS1 Leishmania with a size of 343 bp in all tubes was noticed, while for SF just the first two dilution SF primer tubes (0.8 and 0.08 uM) showed bands with 217 bp molecular weight. For that, dilution 0.08 μ M as a limit factor for 18S rRNA gene was selected as final concentration for 18S rRNA primer in each reaction should be 0.08 μ M to give stable amplification product in the multiplex PCR reaction, see figure (3.5).

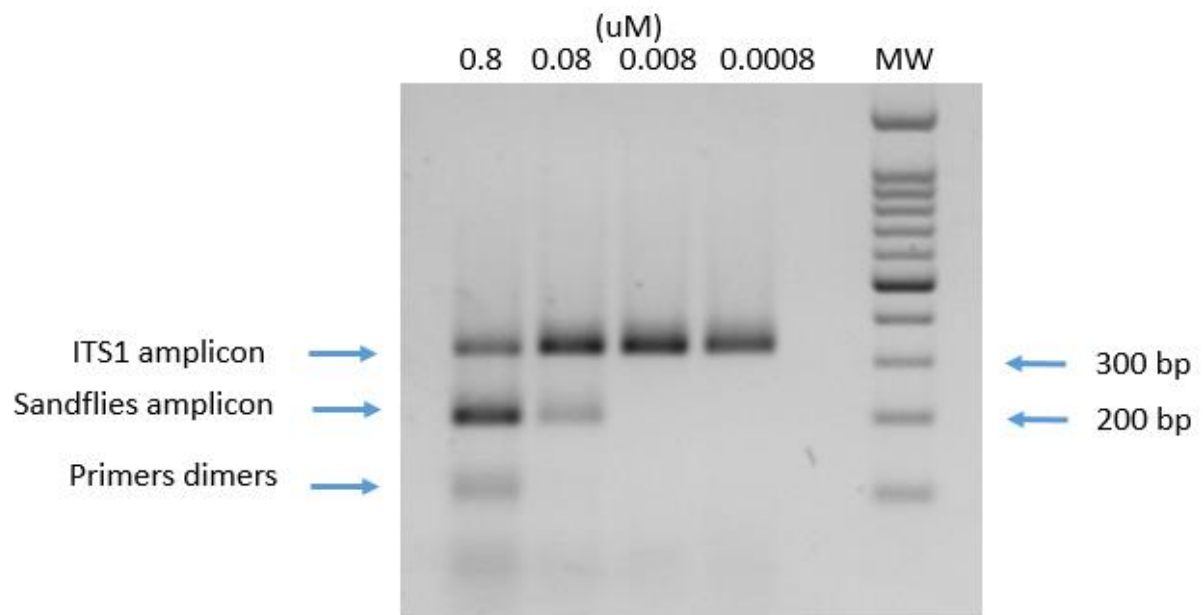


Figure 3.5: multiplex PCR for ITS1 and 18S rRNA primers dilutions from 0.8-0.0008 μM with constant *Leishmania* ITS1 primers concentrations (1 μM).

3.2 Optimization with different global worldwide *Leishmania* species

Nine different global *Leishmania* species were used and then measure their DNA concentration then run ITS1 PCR on these samples they gave clear bands for ITS1 gene, figure (3.6).

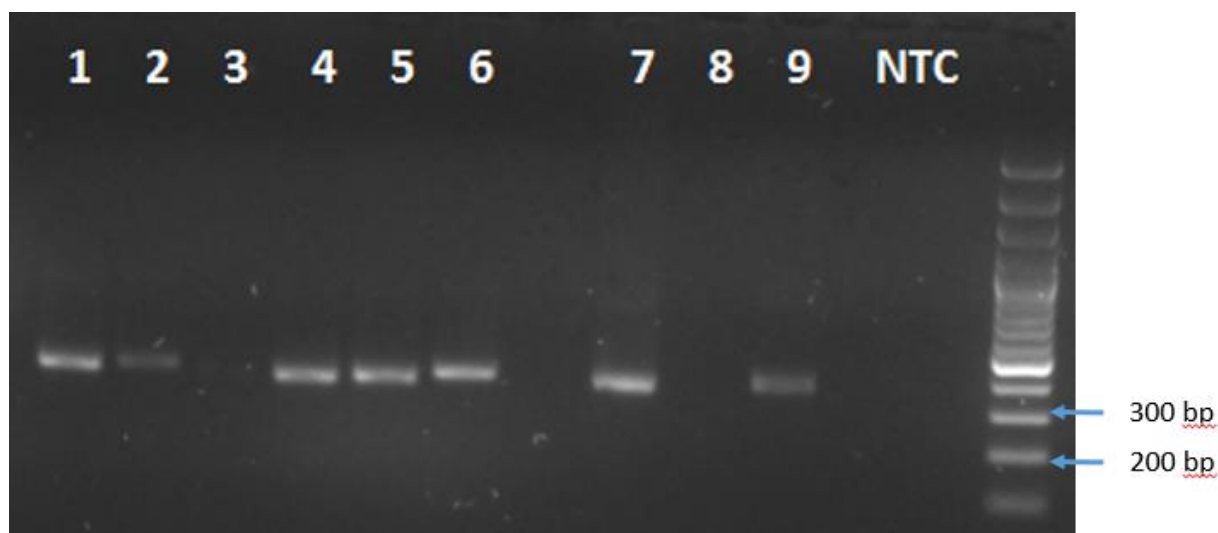


Figure 3.6: ITS1 result of different *Leishmania* samples used as in appendix 3. NTC: negative control, (1-9) number of samples. **1:** *L. tropica* **2:** *L. aethiopica* **3:** *L. infantum* **4:** *L. infantum* **5:** *L. donovani* **6:** *L. major* **7:** *L. donovani* **8:** *L. tropica* **9:** *L. tropica*.

3.3 Species specific probes design

PCR purification and barcode Indices PCR was done for NGS analysis using specific probes for all *Leishmania* species (GATCATTTTCCGATG) and we get the result as shown in table (3.2).

Nine samples gave ITS1 sequences were obtained and significant reads counts were seen. These sequences were compared by multiple-alignment of all species and for the same species (obtained from the NCBI database) to find different genotype see Table (3.2). Multiple alignment for all *leishmania* samples was done to find specific probes for each species and genotype. These probes were used in building the workflow. Figure (3.7) showed DNA sequence pattern for *Leishmania* species, these pattern are shown into different colored boxes upon the DNA sequence pattern, like green box for *L. donovani* and *L. infantum*, but in the red box, specific probe for *L. donovani* was shown, orange box for *L. major*. Specific probes to differentiate between *L. tropica* G1, *L. tropica* G2 and *L. aethiopica* depend on DNA polymorphism in the pink box. Other black and blue boxes show another polymorphisms and deletion in DNA sequence for different *Leishmania* species.

Table 3.2: NGS analysis result for eight Leishmania samples.

	Leishmania species	Total reads	Leishmania sequence
1	<i>L. tropica</i> G1	35628	>1-21616 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAAAACCTCGGGGAG GCCTATATATTATACATTATATATAGGCCTTTCCACACATACACAGCAAACCTTTATA CTCGAAGTTTGCAGTAAACAAAAGGCCGATCGA
2	<i>L. aethiopica</i>	32285	>1-15944 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATACAAAACCTCGGGCAGGC CTATTATATATATTATTATAGGCCTTTCCACATACACAGCAAACCTTTTATACTCG AAGTTTGCAGTAAAGAAAAGGCCGATCGACGTTAT
3	<i>L. infantum</i>	6656	>1-5855 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAACCTCGGGGAGACCTA TGTATATATATGTAGGCCTTTCCACATACACAGCAAAGTTTGTACTCAAAATTTGCA GTAAAAAAAAGGCCGATCGACGTTATAACGCA
4	<i>L. infantum</i>	32017	>1-29488 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAACCTCGGGGAGACC TATGTATATATATGTAGGCCTTTCCACATACACAGCAAAGTTTGTACTCAAAATT TGCAGTAAAAAAAAGGCCGATCGACGTTATAACGCA
5	<i>L. donovani</i>	16581	>1-8559 AGCTGGATCATTTTCCGATGATTACACCAAAAAACATATACAACCTCGGGGAG ACCTATGTATATATATATGTAGGCCTTTCCACATACACAGCAAAGTTTGTACTCA AAATTTGCAGTAAAAAAAAGGCCGATCGACGTTATAA
6	<i>L. major</i>	11626	>1-6517 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAACCTCGGGGAGGCT TATTCTATATATATATAGTATAGGCCTTTCCACATACACAGCAAACCTTTTATACTCA AAATTTGCAGTAAAAAAAAGGCCGATCGACGTTGTA
7	<i>L. tropica</i> G1 <i>L. donovani</i>	11508 15781	>1-6992 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAAAACCTCGGGG AGGCCTATATATTATACATTATATATAGGCCTTTCCACACATACACAGCAAACCTTT TATACTCGAAGTTTGCAGTAAACAAAAGGCCGATCGA >1-7766 AGCTGGATCATTTTCCGATGATTACACCAAAAAACATATACAACCTCGGGGAGA CCTATGTATATATATATGTAGGCCTTTCCACATACACAGCAAAGTTTGTACTCAA AATTTGCAGTAAAAAAAAGGCCGATCGACGTTATAA
8	<i>L. tropica</i> G2	3664	>1-2075 AGCTGGATCATTTTCCGATGATTACACCCCAAAAAACATATACAAAACCTCGGGGA

			GGCCTATATATTATACATTATATATAGGCCTTTCCACACATACACAGCAAACCTTTTA TACTCGAAGTTTGCAGTAAACAAAAGGCCGATCGA
9	<i>L. tropica</i> G1	11861	>1-6895 AGCTGGATCATTTCGATGATTACACCCCAAAAAACATATACAAAACCTCGGGGA GGCCTATATATTATACATTATATATAGGCCTTTCCACACATACACAGCAAACCTTTTA TACTCGAAGTTTGCAGTAAACAAAAGGCCGATCGA

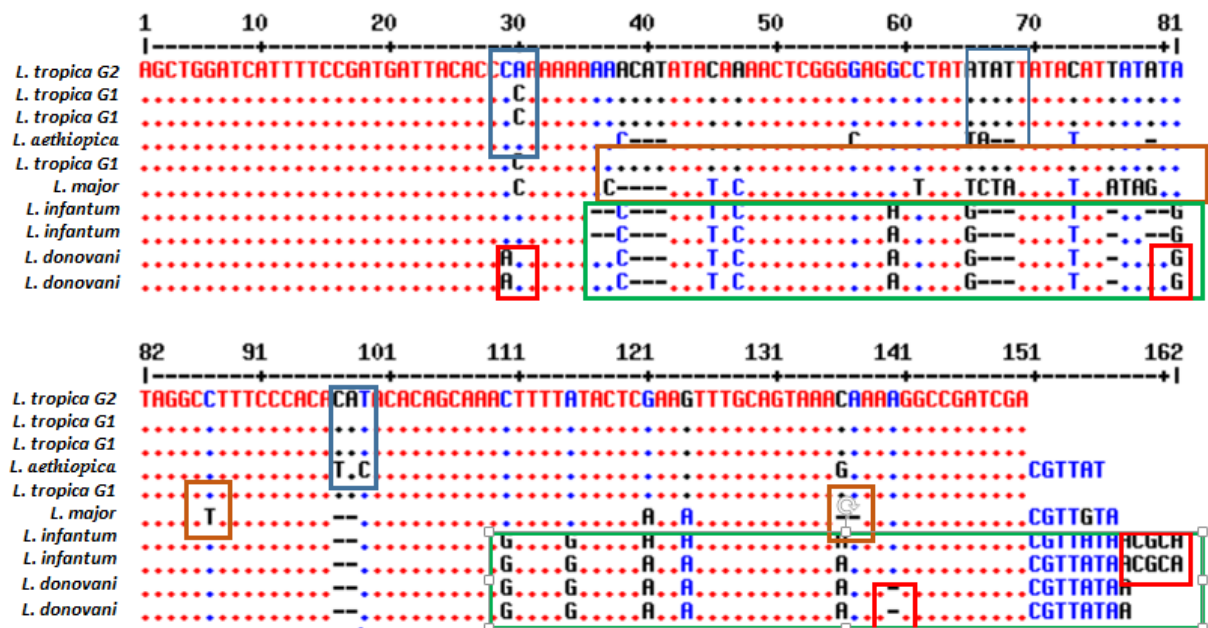


Figure 3.7: DNA sequencing for different *Leishmania* species identification probes.

3.4 Detection of minimal *Leishmania* ITS-PCR copy number

In the next multiplex PCR, constant DNA concentration for sand fly *Ph. serganti* and serial dilution with dilution factor (1:10) from *L. tropica* DNA, and we found the constant copy number for sand flies in all samples and the decrease in the copy number of *Leishmania* compatible with their serial dilution as in table (3.3).

Table 3.3: NGS analysis result for Multiplex PCR with serial dilution of *L. tropica*.

Sample Number	Sand fly reads number	Leishmania reads number
1	106957	19617
2	110496	17037
3	129792	5770
4	175327	3262
5	192389	493
6	206048	0
7	199748	0

3.5 Field Samples

Field caught samples contains 171 sand flies collected in three traps catches, the result of Multiplex PCR showed in figure (3.8), the negative control (NTC) showed no bands and there are clear bands for sand fly PCR (1-60) and some samples showed clear band for infection with *Leishmania* (9, 17, 21, 36, 51, and 55).

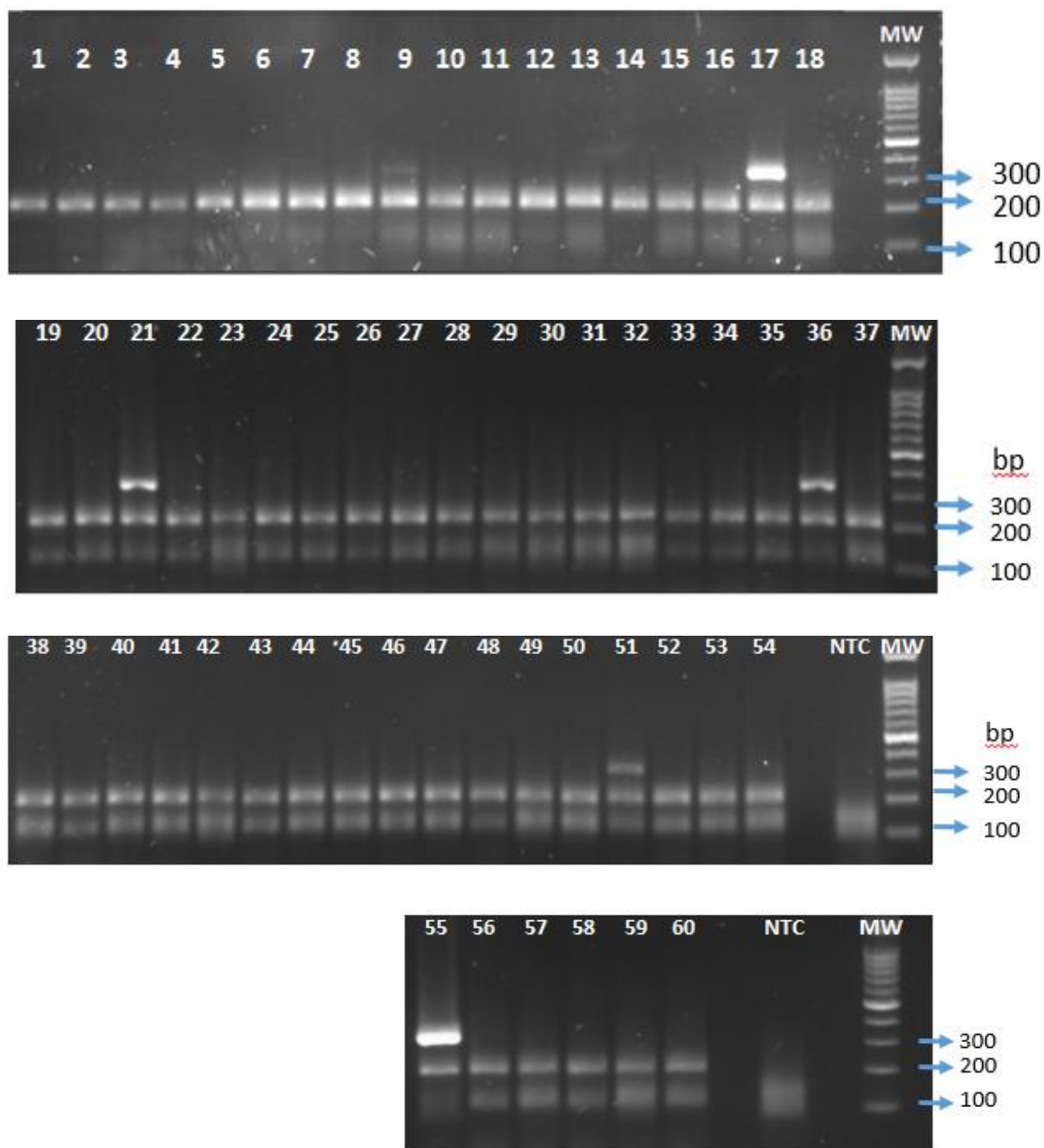


Figure 3.8 Multiplex PCR for ITS1 and 18S rRNA for 60 sand flies as representative for the total DNA samples, with primer dimer band size 130 bp .

3.6 NGS data and statistical analysis

NGS analysis results using specific probes for *Leishmania* and sand fly are shown in the appendix (4). The most detected sand fly species was *Ph. sergenti* (86%, 147/171), *Ph. syriacus* (3.5%, 6/171), *S. dentata* (3.5%, 6/171), *Ph. perfiliewi* (1.8%, 3/171), *Ph. papatasi* (1.8%, 3/171), *S. dubia* (1.2%, 2/171), *Ph. tobbi* (1.2%, 2/171), *S. ghesquierei* (0.6%, 1/171) and *S. schwetzi* (0.6%, 1/171) see table (3.4). Result for sand fly identification by NGS and the traditional method by dissection were compared as shown in table (3.5) and table (3.6), data agreement using the statistical significance of NGS results in p value < 0.05.

Table 3.4: Frequency and percentage of sand fly species identification by NGS.

		Frequency	Percent %
SF species	<i>Ph. sergenti</i>	147	86.0
	<i>S. dentata</i>	6	3.5
	<i>Ph. syriacus</i>	6	3.5
	<i>Ph. perfiliewi</i>	3	1.8
	<i>Ph. papatasi</i>	3	1.8
	<i>Ph. tobbi</i>	2	1.2
	<i>S. dubia</i>	2	1.2
	<i>S. ghesquierei</i>	1	0.6
	<i>S. schwetzi</i>	1	0.6
	Total	171	100 %

Table 3.5: Cross tabulation between NGS method for Phlebotomus genera and traditional method result.

		Sand fly species by NGS					Total
		<i>Ph. papatasi</i>	<i>Ph. perfiliewi</i>	<i>Ph. sergenti</i>	<i>Ph. syriacus</i>	<i>Ph. tobbi</i>	
SF species by dissection	<i>Ph. kazeruni</i>	0	0	1	0	0	1
	<i>Ph. papatasi</i>	3	0	0	0	0	3
	<i>Ph. perfiliewi</i>	0	3	0	3	0	6
	<i>Ph. sergenti</i>	0	0	144	0	0	144
	<i>Ph. tobbi</i>	0	0	0	3	2	5
	<i>S. theodoi</i>	0	0	1	0	0	1
	<i>S. tiberiadis</i>	0	0	1	0	0	1
	Total	3	3	147	6	2	161

Table 3.6: Cross tabulation between NGS method for Sergentomyia genera and traditional method result.

		Sand fly species by NGS				Total
		<i>S. dentata</i>	<i>S. dubia</i>	<i>S. ghesquierei</i>	<i>S. schwetzi</i>	
SF species by dissection	<i>Ph. kazeruni</i>	1	0	0	0	1
	<i>S. dentata</i>	2	0	0	0	2
	<i>S. dubia</i>	0	2	0	0	2
	<i>S. fallax</i>	3	0	0	0	3
	<i>S. schwetzi</i>	0	0	0	1	1
	<i>S. tiberiadis</i>	0	0	1	0	1
Total		6	2	1	1	10

For Leishmania, the analysis done by using specific forward and reverse probes for Leishmania to give an accurate read number. The cutoff value of 200 reads per sample was

setup to determine *Leishmania* positivity (Early et al., 2018). As shown 8/171(4.7%) samples were positives for *Leishmania*, and 163/171(95.3%) samples were negatives for *Leishmania* as following: 137/163(84.0%) give zero reads per sample, 26/163(16.0%) sample give low numbers of *Leishmania* sequences between 1-200 reads per sample, see table (3.7).

Table 3.7: Frequency of sand flies that have leishmania parasite (negative < 200 copy number of leishmania parasite) and (Positive > 200 copy number of leishmania parasite).

		SF Frequency	Percent %
<i>Leishmania</i>	Negative	163	95.3
	Positive	8	4.7
	Total	171	100 %

A cross-tabulation comparison between ITS1 PCR results for 129 sand fly samples that done by Leishmaniasis Research Unit, Jericho, Palestine (LRU) and their infection rate was (22/129)17%, and our result by NGS give infection rate (5/129) 3.9%. Statistical significance correlation using two-tailed Fisher's exact test between the results showed p value < 0.0001, which considered to be extremely statistically significant, see table (3.8).

Table 3.8: Cross tabulation between NGS and traditional method result for Leishmania (negative < 200 copy number of leishmania parasite) and (Positive > 200 copy number of leishmania parasite).

		NGS method		Total
		Negative	Positive	
ITS1 PCR method	Negative	107	0	107
	Positive	17	5	22
Total		124	5	129

Chapter Four:

4. Discussion

This study is an important study takes place in worldwide and in Palestine that optimized novel method in using NGS in identification of sand flies (*Phlebotomus* and *Sergentomyia* species) and their population abundance. In addition, detection and identification of leishmania species and some specific genotypes, and parasite load rates in these sandflies in one tube experiment. The method will have a significant advantage over the classical molecular methods used to perform high-throughput sequencing in ecological samples.

Identification of the *Leishmania* and sand fly species, especially in endemic areas is important for different reasons like appropriate treatment and estimation of the patient's prognosis. Sand flies consist of more than 500 species, only a few species are medically important (Kato et al., 2005). For many years and until now many of the studies used the traditional method of sand fly species identification by using morphological taxonomy of the phlebotomine Sandflies, mainly by mounting of the head, the abdomen and their internal structures, such as spermatheca, cibarium, and pharynx in females and terminal genitalia in males under stereomicroscope (Giantsis et al., 2017). This method has many disadvantages, for example, requires expert entomologists to differentiate between the morphology of sand fly species to minimize the error of classification. Also, this method

needs special storage conditions for samples and time consuming when dealing with large sample size (Tiwary, Kumar, Rai, & Sundar, 2014).

In recent years, many studies use a molecular techniques to differentiate between sand fly species, like differentiate only *Ph. argentipes* from other sand flies (Surendran, Karunaratne, Adamsn, Hemingway, & Hawkes, 2005), or differentiate between only two species *Ph. argentipes* and *Ph. papatasi* (Manonmani, Mathivanan, Srinivasan, & Jambulingam, 2010). In this study using NGS technique we can differentiate between more than 15 species of sand fly and detect Leishmania DNA within sand flies.

Studying sand fly and leishmania is important for Leishmaniasis disease control. However, most of the currently available methods of identification of sand fly and leishmania relatively low throughput and cannot be applied to a large sample size. The developed high-throughput screening is cost-effective assay for sand fly identification and detecting and identification of Leishmania species.

High-throughput screening (HTS) is one of the modern tools used in drug design and applied for biological and chemical sciences (Szymański, Markowicz, & Mikiciuk-Olasik, 2012). HTS has many molecular advantages like testing hundreds or thousands of samples simultaneously, ability to analyze multiple genetic loci simultaneously for a single sample, and get results within a short time (Wiita & Schrijver, 2011). These advantages were shown in this research will change and make the method of identification of sand fly species and leishmania abundances more sensitive, specific and easy with very rapid result especially when we took about a huge number of samples, and finally with cost effective system. To get high sensitivity and specify we design PCR primer for ITS1 gene for identification of Leishmania parasite cause by using this gene we get the most accurate method to detect leishmania with higher sensitivity and specify (Hitakarun et al., 2014).

There are differences in sand fly species identification between classical methods using dissection and our NGS method. Some of these sand flies were identified as *Ph. paptasi* by classical method then identified as *Ph. sergenti* by NGS. After that, the entomologist reviewed the samples and found were incorrect in classification and confirm that's *Ph. sergenti* as NGS result. Another sample was classified as *S. christophersi* and by NGS classified as *S. schwetzi*. The entomologist reviewed the samples then confirm the NGS result, and retreat on the previous result.

Differences in the result can be explained as the following. Some of these sand flies were identified as *Ph. tobbi* by dissection and identified to be *Ph. syriacus* by using NGS method. These two species belong to the subgenus *Larroussius* and it is very difficult to differentiate between the species of this subgenus morphologically. The only way to do is by distinguishing spermathecal duct which mostly not clear. Unfortunately, the spermatheca in this sample was lost during sample preparation and the head also lost during reexamination. In this study, results showed the quality and preservation of samples is not required or crucial as the traditional classical method for taxonomy identification. Samples from field mainly got exposed to a lot of harsh conditions during collection and treatment mainly the delicate parts of the sandflies.

Other samples were identified as *S. fallax* by dissection and identified as *S. dentate* by using NGS method. As an entomologist explained regarding the Pharynx shape of sand fly the possibility to be *S. fallax* is 70% according to most references, but it may be *S. dentate* with possibility of 30% as we find by using NGS method, also the two species belong to the same subgenus *Sergentomyia* and some morphological features nearly comparable. This made the ability to differentiate between two these species in traditional method difficult if not impossible. Some samples identified as *S. dubia* and as *S. theodori* in dissection method. After review the entomologist they found the possibility to be *S. dubia* is 50% and this species was not reported in Palestine before. Two samples identified as *S. ghesquierei* and *Ph. sergenti* by using NGS method, but by a classical method the two sample identify as *S. tiberiadis* with possibility 95%. However, P-value for our result was < 0.01 which mean that our result statistically significant.

Low numbers of Leishmanial sequences between 1-200 reads per sample were obtained from some samples and were considered as negatives, they were considered as technical contaminant due to lack of ability to determine the species after looking for similarities in blast analysis. However, contaminating sequences were removed from only 16% of samples and occurred with low read counts (typically in the tens of reads). A related issue pertains to the possibility of cross-contamination between samples as stated previously (Salipante et al., 2013).

Also, the abundance of leishmania parasite in all 171 sand flies was 8/171 (4.7%) carried Leishmania DNA, all of them were of *L. tropica* species, as expected since Tubas endemic region for CL caused by *L. tropica* parasites. ITS1 PCR done for 129 of these samples in

LRU and the Leishmania abundance was 22/129 (17%), but by our method using NGS gives 5/129 (3.9%). This study result is reasonable when compared to different global studies, like in Brazil the infection rate was 1.56 % (Felipe et al., 2011). USA the infection rate was detected in the winter season with 2.84% (Tiwary et al., 2013), in Iran was 4.7%. Our result by NGS is reasonable and close to infection rate in different countries also NGS result's were confirmed than the ITS1 assay because DNA sequences for each Leishmania parasite were verified. While in the traditional ITS1 PCR Sanger sequencing method failed to have DNA sequence of more than 90% of the positive samples, which propose the theory of cross contamination within the traditional PCR so these samples could most probably be a false positive.

In this study, we recorded nine different sand fly species in Palestine regions (*Ph. sergenti*, *S. dentate*, *S. dubia* *Ph. syriacus*, *S. ghesquierei*, *S. schwetzi*, *Ph. papatasi*, *Ph. perfiliewi*, and *Ph. tobbi*). *Ph. sergenti* was the most abundant *Phlebotomus* species as in Jenin District (Sawalha et al., 2017). In other country they found that *Ph. papatasi* was the most abundant in southern Jordan Valley in 1992 (Janini, Saliba, & Kamhawi, 1995). This species has a wide distribution in the Middle East and has been confirmed as the vector of CL caused by *L. tropica* (Es-Sette et al., 2014). Most of the studies highlighted only *Phlebotomus* species sand flies role in the leishmania transmission cycle, but in a recent study in Tunisia, they found that *S. minuta* have *L. major* DNA which causes CL. So studying *Sergentomyia* species is important to explain the role of this species in the transmission of Leishmania parasites in different endemic regions (Jaouadi et al., 2015)

Furthermore, the NGS system showed the ability to identify sandflies from Latin America *Lutzomyia spp.* Which is a vector of Leishmaniasis in Latin America, and *Ph. argentipes* as a vector for VL in India, so the system proved its ability to identify worldwide sandflies and not restricted to the Middle East sand fly fauna.

In conclusion:

This is an important successful study using NGS to detect parasite, vector and identify their species by experimental and natural infections, the study considered a millstone in further studies of vector–parasite co evolution. The method is able to perform high-throughput sequencing in ecological samples. These results will help in detecting the transmission of several potential vectors that vary in their spatial and geographical

distribution, which could explain the high prevalence of Leishmaniasis cases in specific endemic regions. The study could be applied on new world leishmaniasis, since here, vector of Leishmaniosis (*Lutzumia* species) in South America was used as a reference.

Study limitation

This new high throughput system requested a large number of samples, since using a low sample number made it expensive. NGS machines not available in most of scientific lab or Universities in Palestine for the time being, otherwise send samples to outsources increase costs, requested high experience to deal with NGS machine and result analysis.

Recommendations

This study could carry important recommendations for World Health Organization to use this high throughput system in endemic areas that suffering from leishmaniasis disease especially in Africa and India. It highlights the importance and the prognostic value of using NGS technique as an alternative method for traditional method.

REFERENCES

- Al-Jawabreh, A., Dumaidi, K., Ereqat, S., Al-Jawabreh, H., Nasereddin, A., Azmi, K., . . . Abdeen, Z. (2017). Molecular epidemiology of human cutaneous leishmaniasis in Jericho and its vicinity in Palestine from 1994 to 2015. *Infection, Genetics and Evolution*, 50, 95-101.
- Alemayehu, B., & Alemayehu, M. (2017). Leishmaniasis: a review on parasite, vector and reservoir host. *Health Science Journal*, 11(4), 1.
- Alexander, B., & Maroli, M. (2003). Control of phlebotomine sandflies. *Medical and veterinary entomology*, 17(1), 1-18.
- Alvar, J., Velez, I. D., Bern, C., Herrero, M., Desjeux, P., Cano, J., . . . Team, W. L. C. (2012). Leishmaniasis worldwide and global estimates of its incidence. *PloS one*, 7(5), e35671.
- Basho, R. K., & Eterovic, A. K. (2015). Clinical applications and limitations of next-generation sequencing. *American Journal of Hematology/Oncology®*, 11(3).
- Bates, P. A. (2008). Leishmania sand fly interaction: progress and challenges. *Current opinion in microbiology*, 11(4), 340-344.
- Berman, J. (2006). Visceral leishmaniasis in the New World & Africa. *Indian Journal of Medical Research*, 123(3), 289.
- Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A., & the Galaxy, T. (2010). Manipulation of FASTQ data with Galaxy. *Bioinformatics*, 26(14), 1783-1785. doi:10.1093/bioinformatics/btq281
- Bounamous, A., Lehrter, V., Hadj-Henni, L., Delecolle, J.-C., & Depaquit, J. (2014). Limits of a rapid identification of common Mediterranean sandflies using polymerase chain reaction-restriction fragment length polymorphism. *Memórias do Instituto Oswaldo Cruz*, 109(4), 466-472.
- Casari, A. E., de Oliveira, L. P., Alonso, D. P., de Oliveira, E. F., Barrios, S. P. G., Infran, J. d. O. M., . . . Ribolla, P. E. M. (2017). Standardization of DNA extraction from sand flies: Application to genotyping by next generation sequencing. *Experimental parasitology*, 177, 66-72.
- Chemkhi, J., Guerbouj, S., Saadawi, W., Shaibi, T., Faris, S., Ghawar, W., . . . Ben Salah, A. (2018). Presence of *Sergentomyia* (*Parrotomyia*) *lewisii* (Diptera: Psychodidae) in Tunisia. *Journal of medical entomology*, 56(2), 560-564.
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic acids research*, 16(22), 10881-10890.

- Dávila, A., & Momen, H. (2000). Internal-transcribed-spacer (ITS) sequences used to explore phylogenetic relationships within *Leishmania*. *Annals of Tropical Medicine & Parasitology*, 94(6), 651-654.
- Desjeux, P. (2001). Worldwide increasing risk factors for leishmaniasis. *Medical microbiology and immunology*, 190(1-2), 77-79.
- Dostálová, A., & Volf, P. (2012). *Leishmania* development in sand flies: parasite-vector interactions overview. *Parasites & vectors*, 5(1), 276.
- Early, A. M., Daniels, R. F., Farrell, T. M., Volkman, S. K., Wirth, D. F., MacInnis, B. L., & Neafsey, D. E. (2018). Amplicon deep sequencing of low-density *Plasmodium falciparum* infections: an evaluation of analysis approaches. *bioRxiv*, 453472.
- Es-Sette, N., Ajaoud, M., Bichaud, L., Hamdi, S., Mellouki, F., Charrel, R. N., & Lemrani, M. (2014). *Phlebotomus sergenti* a common vector of *Leishmania tropica* and Toscana virus in Morocco. *Journal of vector borne diseases*, 51(2), 86.
- Felipe, I. M. A., Aquino, D. M. C. d., Kuppinger, O., Santos, M. D. C., Rangel, M. E. S., Barbosa, D. S., . . . Caldas, A. d. J. M. (2011). *Leishmania* infection in humans, dogs and sandflies in a visceral leishmaniasis endemic area in Maranhão, Brazil. *Memórias do Instituto Oswaldo Cruz*, 106(2), 207-211.
- Fokialakis, N., Kalpoutzakis, E., Tekwani, B., Khan, S., Kobaisy, M., Skaltsounis, A., & Duke, S. (2007). Evaluation of the antimalarial and antileishmanial activity of plants from the Greek island of Crete. *Journal of natural medicines*, 61(1), 38-45.
- Giantsis, I. A., Chaskopoulou, A., & Claude Bon, M. (2017). Direct multiplex PCR (dmPCR) for the identification of six phlebotomine sand fly species (Diptera: Psychodidae), including major leishmania vectors of the Mediterranean. *Journal of economic entomology*, 110(1), 245-249.
- Hitakarun, A., Tan-Ariya, P., Siripattanapipong, S., Mungthin, M., Piyaraj, P., Naaglor, T., . . . Leelayoova, S. (2014). Comparison of PCR methods for detection of *Leishmania siamensis* infection. *Parasites & vectors*, 7(1), 458.
- Janini, R., Saliba, E., & Kamhawi, S. (1995). Species composition of sand flies and population dynamics of *Phlebotomus papatasi* (Diptera: Psychodidae) in the southern Jordan Valley, an endemic focus of cutaneous leishmaniasis. *Journal of medical entomology*, 32(6), 822-826.
- Jaouadi, K., Ghawar, W., Salem, S., Gharbi, M., Bettaieb, J., Yazidi, R., . . . Salah, A. B. (2015). First report of naturally infected *Sergentomyia minuta* with *Leishmania major* in Tunisia. *Parasites & vectors*, 8(1), 649.
- Kakarsulemankhel, J. K. (2010). Taxonomic review of sand flies of the subgenus *Paraphlebotomus* Theodor (Diptera: Psychodidae). *Pakistan Entomologist*, 32(2), 125-147.

- Kassem, H. A., Fryauff, D. J., El Sawaf, B. M., Shehata, M. G., & Shoumar, N. F. (1990). Electrophoretic comparison of the leishmania vectors *Phlebotomus papatasi* and *P. Langeroni* (diptera: psychodidae). *Journal of medical entomology*, 27(4), 592-601.
- Kato, H., Uezato, H., Katakura, K., Calvopina, M., Marco, J. D., Barroso, P. A., . . . Iwata, H. (2005). Detection and identification of *Leishmania* species within naturally infected sand flies in the Andean areas of Ecuador by a polymerase chain reaction. *The American journal of tropical medicine and hygiene*, 72(1), 87-93.
- Khraiwesh, M., Leed, S., Roncal, N., Johnson, J., Sciotti, R., Smith, P., . . . Hickman, M. (2016). Antileishmanial activity of compounds derived from the medicines for malaria venture open access box against intracellular *Leishmania major* amastigotes. *The American journal of tropical medicine and hygiene*, 94(2), 340-347.
- Killick-Kendrick, R. (1999). The biology and control of phlebotomine sand flies. *Clinics in dermatology*, 17(3), 279-289.
- Lane, R. P. (1986). *sandflies of Egypt (Diptera, Phlebotominae)*: British Museum (Natural History).
- Lewis, D. (1971). Phlebotomid sandflies. *Bulletin of the World Health Organization*, 44(4), 535.
- Lewis, D. J. (1982). A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae).
- Manonmani, A. M., Mathivanan, A., Srinivasan, R., & Jambulingam, P. (2010). Species-diagnostic polymerase chain reaction assays for *Phlebotomus argentipes* and *Phlebotomus papatasi*, Vectors of *Leishmania*. *Journal of medical entomology*, 47(5), 743-747.
- Monroy-Ostria, A., Nasereddin, A., Monteon, V. M., Guzmán-Bracho, C., & Jaffe, C. L. (2014). ITS1 PCR-RFLP diagnosis and characterization of *Leishmania* in clinical samples and strains from cases of human cutaneous leishmaniasis in states of the Mexican Southeast. *Interdisciplinary perspectives on infectious diseases*, 2014.
- Moore, J., Kelly, T., Killick- Kendrick, R., Killick- Kendrick, M., Wallbanks, K., & Molyneux, D. (1987). Honeydew sugars in wild- caught *Phlebotomus ariasi* detected by high performance liquid chromatography (HPLC) and gas chromatography (GC). *Medical and veterinary entomology*, 1(4), 427-434.
- Organization, W. H. (2018, March 14). Leishmaniasis. Retrieved from <http://www.who.int/en/news-room/fact-sheets/detail/leishmaniasis>
- Ready, P. D., Smith, D. F., & Killick-Kendrick, R. (1988). DNA hybridizations on squash-blotted sandflies to identify both *Phlebotomus papatasi* and infecting *Leishmania major*. *Med Vet Entomol*, 2(2), 109-116.
- Rocha, V. P. C., Nonato, F. R., Guimarães, E. T., de Freitas, L. A. n. R., & Soares, M. B. P. (2013). Activity of antimalarial drugs in vitro and in a murine model of cutaneous leishmaniasis. *Journal of medical microbiology*, 62(7), 1001-1010.

- Salam, N., Al-Shaqha, W. M., & Azzi, A. (2014). Leishmaniasis in the Middle East: incidence and epidemiology. *PLoS neglected tropical diseases*, 8(10), e3208.
- Salipante, S. J., Sengupta, D. J., Rosenthal, C., Costa, G., Spangler, J., Sims, E. H., . . . Cookson, B. T. (2013). Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. *PloS one*, 8(5), e65226.
- Sawalha, S. S., Ramlawi, A., Sansur, R. M., Salem, I. M., & Amr, Z. S. (2017). Diversity, ecology, and seasonality of sand flies (Diptera: Psychodidae) of the Jenin District (Palestinian Territories). *Journal of Vector Ecology*, 42(1), 120-129.
- Sawalha, S. S., Shtayeh, M. S., Khanfar, H. M., Warburg, A., & Abdeen, Z. A. (2003). Phlebotomine sand flies (Diptera: Psychodidae) of the Palestinian West Bank: potential vectors of leishmaniasis. *Journal of medical entomology*, 40(3), 321-328.
- Schönian, G., Nasereddin, A., Dinse, N., Schweynoch, C., Schallig, H. D., Presber, W., & Jaffe, C. L. (2003). PCR diagnosis and characterization of *Leishmania* in local and imported clinical samples. *Diagnostic microbiology and infectious disease*, 47(1), 349-358.
- Steverding, D. (2017). The history of leishmaniasis. *Parasites & vectors*, 10(1), 82.
- Sundar, S., Lockwood, D. N., Agrawal, G., Rai, M., Makharia, M., & Murray, H. W. (2001). Treatment of Indian visceral leishmaniasis with single or daily infusions of low dose liposomal amphotericin B: randomised trialCommentary: cost and resistance remain issues. *Bmj*, 323(7310), 419-422.
- Sundar, S., & Rai, M. (2002). Laboratory diagnosis of visceral leishmaniasis. *Clinical and diagnostic laboratory immunology*, 9(5), 951-958.
- Surendran, S., Karunaratne, S., Adamsn, Z., Hemingway, J., & Hawkes, N. (2005). Molecular and biochemical characterization of a sand fly population from Sri Lanka: evidence for insecticide resistance due to altered esterases and insensitive acetylcholinesterase. *Bulletin of entomological research*, 95(4), 371-380.
- Szymański, P., Markowicz, M., & Mikiciuk-Olasik, E. (2012). Adaptation of high-throughput screening in drug discovery—toxicological screening tests. *International journal of molecular sciences*, 13(1), 427-452.
- Talmi-Frank, D., Jaffe, C. L., Nasereddin, A., Warburg, A., King, R., Svobodova, M., . . . Baneth, G. (2010). *Leishmania tropica* in rock hyraxes (*Procavia capensis*) in a focus of human cutaneous leishmaniasis. *The American journal of tropical medicine and hygiene*, 82(5), 814-818.
- Tiwary, P., Kumar, D., Mishra, M., Singh, R. P., Rai, M., & Sundar, S. (2013). Seasonal variation in the prevalence of sand flies infected with *Leishmania donovani*. *PloS one*, 8(4), e61370.
- Tiwary, P., Kumar, D., Rai, M., & Sundar, S. (2014). PCR-RFLP based method for molecular differentiation of sand fly species *Phlebotomus argentipes*, *Phlebotomus*

papatasi, and *Sergentomyia babu* found in India. *Journal of medical entomology*, 49(6), 1515-1518.

Wheeler, R. J., Gluenz, E., & Gull, K. (2011). The cell cycle of *Leishmania*: morphogenetic events and their implications for parasite biology. *Molecular microbiology*, 79(3), 647-662.

Wiita, A. P., & Schrijver, I. (2011). Clinical application of high throughput molecular screening techniques for pharmacogenomics. *Pharmacogenomics and personalized medicine*, 4, 109.

Appendix 1: Different reference sand fly species used in the study.

Sample number	Sand fly species	Endemic area	Reference
A	<i>Ph. argentipes</i>	Indian subcontinent, (India, Nepal, Bangladesh)	Umakant Sharma & Sarman Singh et al. 2008
B	<i>Ph. syriacus</i>	Europe, the Middle East and North Africa	Maroli et al. 2012
C	<i>Ph. perfelewi</i>	Mediterranean basin and Central Asia	Depaquit et al. 2013
D	<i>Ph. tobbi</i>	Middle East, Eroupe and North Africa	Maroli et al. 2012
E	<i>Ph. papatasi</i>	North Africa, central and west Asia	Umakant Sharma & Sarman Singh et al. 2008
F	<i>Ph. sergenti</i>	Central and west Asia and western India	Umakant Sharma & Sarman Singh et al. 2008)
G	<i>Sergentomyia</i> genus	Europe, the Middle East and North Africa	(Chemkhi et al., 2018)
H	<i>Lutzomyia</i> spp	South America	Umakant Sharma & Sarman Singh et al. 2008

Appendix 2: Different reference sand fly species with their gene bank accession numbers used in the study.

Sample number	Gene bank number	Sand fly species
1	AJ244391	<i>Ph. perfiliewi</i>
2	AJ244391	<i>Ph. perfiliewi</i>
3	AJ244360	<i>Ph. argentipes</i>
4	AJ244376	<i>Ph. syriacus</i>
5	JQ790519	<i>Ph. major</i>
6	JQ790519	<i>Ph. major</i>
7	AJ244384	<i>Ph. tobbi</i>
8	AJ244384	<i>Ph. tobbi</i>
9	MF631017	<i>Ph. papatasi</i>
10	MF631017	<i>Ph. papatasi</i>
11	AJ391727	<i>Ph. sergenti</i>
12	AJ244402	<i>Ph. sergenti</i>
13	AJ391740	<i>S. minuta</i>
14	JQ790518	<i>S. barraudi</i>
15	AJ391737	<i>S. buxtoni</i>
16	kp702939	<i>Ph. lutzomyia</i>

Appendix 3 : Different Reference leishmania species with WHO codes used in the study.

	LRC number	WHO code	Species
1	L1188	MHOM/IL/2005/LRC-L1188	<i>L. tropica</i>
2	L149	MHOM/ET/1972/L102	<i>L.aethiopica</i>
3	L1314	MCAN/IT/2007/LRC-L1314	<i>L. infantum</i>
4	L949	MHOM/IL/2003/LRC-L949	<i>L. major</i>
5	HU3	MHOM/ET/67/HU3	<i>L. donovani</i>
6	5ASKH	MHOM/SU/1973/5ASKH	<i>L. major</i>
7	L1611	MHOM/IL/2012/LRC-L1611	<i>L. tropica</i>
8	L1015	MPRO/IL/2003/HYRAX107	<i>L. tropica</i>
9	L1686	MHOM/IL/2014/LRC-L1686	<i>L. tropica</i>

Appendix 4: NGS analysis result for 173 sand flies samples.

sample number	Sand fly species by dissection	Sand fly species by NGS	Sand fly copy number	ITS1 PCR Leishmania	NGS Leishmania	NGS Leishmania copy number
1	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	37027	Negative	Negative	
2	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	37538	Positive	Negative	
3	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23461	Negative	Negative	
4	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	49697	Negative	Negative	
5	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	39881	Negative	Negative	
6	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	42758	Negative	Negative	
7	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	40900	Negative	Negative	
8	<i>Ph. perfiliewi</i>	<i>Ph. perfiliewi</i>	19319	Negative	Negative	
9	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	34571	Negative	Negative	
10	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	16450	Negative	Negative	
11	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	27753	Positive	Negative	
12	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	33333	Negative	Negative	
13	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	30541	Negative	Negative	
14	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	33884	Negative	Negative	
15	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	29662	Negative	Negative	
16	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	40608	Negative	Negative	
17	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	24851	Negative	Leishmania	1
18	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	42940	Negative	Negative	
19	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	40422	Negative	Negative	
20	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	37963	Negative	Negative	
21	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	40865	Negative	Negative	
22	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	50435	Negative	Negative	
23	<i>Ph. tobbi</i>	<i>Ph. tobbi</i>	25589	Negative	Negative	
24	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	29602	Positive	<i>L. tropica</i> G1	1
25	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	24400	Negative	Negative	
26	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	34542	Negative	Negative	
27	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	24078	Positive	<i>L. tropica</i>	13

					G1	
28	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	21220	Negative	Negative	
29	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23365	Negative	<i>L. tropica</i> G1	1
30	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	37568	Negative	Negative	
31	<i>Ph. perfiliewi</i>	<i>Ph. perfiliewi</i>	19655	Negative	Negative	
32	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	47699	Negative	Negative	
33	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17565	Negative	<i>L. infantum</i>	2
34	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17788	Positive	Negative	
35	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	18123	Positive	<i>L. tropica</i> G1	338
36	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15892	Positive	<i>L. tropica</i> G1	2
37	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	32052	Positive	Negative	
38	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23882	Negative	Negative	
39	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	32296	Negative	Leishmania	1
40	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	44667	Negative	Negative	
41	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	36412	Negative	Negative	
42	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	32566	Negative	Negative	
43	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	42776	Negative	Negative	
44	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	34012	Positive	Negative	
45	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	43685	Negative	Negative	
46	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	22956	Negative	Negative	
47	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	35321	Negative	Negative	
48	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	46633	Negative	Negative	
49	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	35174	Positive	Negative	
50	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17941	Negative	Negative	
51	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	27226	Negative	Negative	
52	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	25399	Positive	<i>L. tropica</i> G1	699
53	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	30272	Negative	Negative	
54	<i>Ph. perfiliewi</i>	<i>Ph. perfiliewi</i>	16569	Negative	Leishmania	2

55	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	28945	Negative	Leishmania	19
56	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	27708	Negative	Negative	
57	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	25194	Negative	Leishmania	1
58	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14755	Negative	Negative	
59	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15522	Negative	Negative	
60	<i>Ph. papatasi</i>	<i>Ph. papatasi</i>	12544	Negative	Negative	
61	<i>Ph. papatasi</i>	<i>Ph. papatasi</i>	10156	Negative	Negative	
62	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	10563	Negative	Negative	
63	<i>Ph. perfiliewi</i>	<i>Ph. syriacus</i>	9163	Negative	Negative	
64	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20963	Negative	Negative	
65	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	24336	Negative	Negative	
66	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14201	Positive	Negative	
67	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17836	Positive	<i>L. tropica</i> G1	266
68	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15117	Negative	Negative	
69	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	22770	Negative	Negative	
70	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19728	Positive	Negative	
71	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19630	Negative	Leishmania	3
72	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19219	Positive	<i>L. tropica</i> G1	12
73	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17371	Negative	Negative	
74	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17936	Negative	Negative	
75	<i>Ph. perfiliewi</i>	<i>Ph. syriacus</i>	17148	Negative	Negative	
76	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	16736	Negative	Leishmania	2
77	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23701	Negative	Negative	
78	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20406	Positive	Negative	
79	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15247	Negative	Negative	
80	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17086	Negative	Negative	
81	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	9919	Negative	Leishmania	5
82	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	21906	Negative	Negative	
83	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14835	Negative	Negative	
84	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	13999	Negative	Negative	

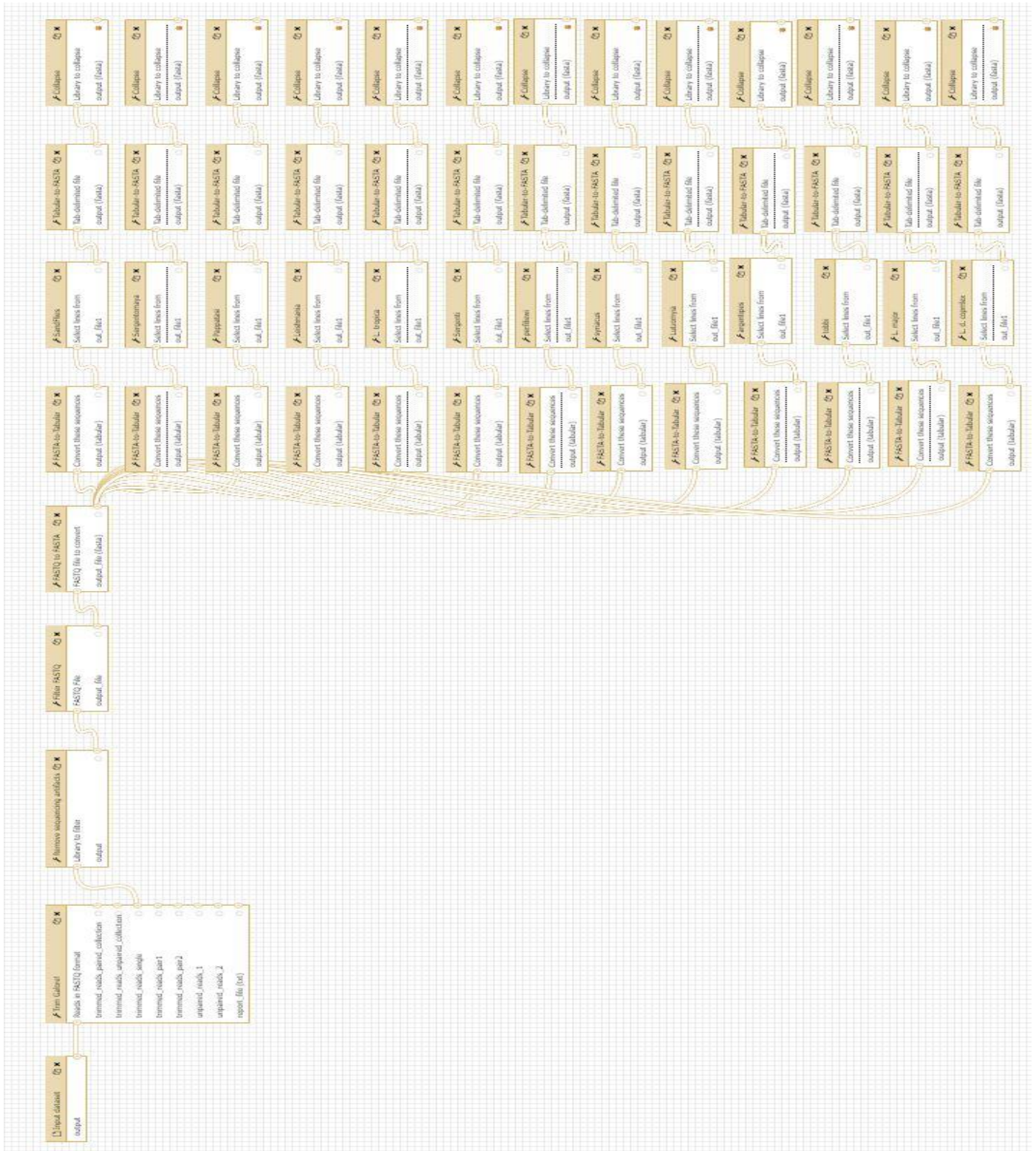
85	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23872	Negative	Negative	
86	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23915	Negative	Negative	
87	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	25581	Negative	Negative	
88	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	29561	Negative	Negative	
89	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	32493	Negative	Negative	
90	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17696	Negative	Negative	
91	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19793	Negative	Negative	
92	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17103	Negative	Negative	
93	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	6532	Negative	Negative	
94	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	18775	Negative	Leishmania	26
95	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20385	Negative	Negative	
96	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	21743	Negative	Negative	
97	<i>Ph. tobbi</i>	<i>Ph. syriacus</i>	19391	Negative	Negative	
98	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14434	Negative	Negative	
99	<i>Ph. kazeruni</i>	<i>Ph. sergenti</i>	16885	Negative	Negative	
100	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19325	Negative	Negative	
101	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	18639	Negative	Negative	
102	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	13533	Negative	Negative	
103	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	260	Negative	Negative	
104	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	7508	Negative	Negative	
105	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23900	Negative	Negative	
106	<i>Ph. tobbi</i>	<i>Ph. syriacus</i>	12712	Negative	Negative	
107	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	18371	Negative	Leishmania	2
108	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	16669	Negative	Negative	
109	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15077	Negative	Negative	
110	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14945	Negative	<i>L. tropica</i> G2	1
111	<i>Ph. perfiliewi</i>	<i>Ph. syriacus</i>	7224	Negative	Leishmania	3

112	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20568	Negative	Negative	
113	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20284	Negative	Negative	
114	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14814	Negative	Negative	
115	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15236	Negative	Negative	
116	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	12455	Negative	Negative	
117	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19500	Negative	Negative	
118	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	8544	Negative	Negative	
119	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	11678	Negative	<i>L. donovani</i>	1
120	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14399	Positive	<i>L. tropica</i> G1	248
121	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	14973	Positive	Negative	
122	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17354	Positive	<i>L. tropica</i> G1	8
123	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15323	Positive	<i>L. tropica</i> G1	10
124	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	17640	Positive	<i>L. tropica</i> G1	1226
125	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	23380	Negative	Negative	
126	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	11153	Negative	Negative	
127	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	19756	Negative	Negative	
128	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	15301	Negative	Negative	
129	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	20920	Positive	Negative	
130	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	94518		Negative	
131	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	213387		<i>L. donovani</i>	142
132	<i>Ph. kazeruni</i>	<i>S. dentata</i>	120709		Negative	
133	<i>S. theodoi</i>	<i>Ph. sergenti</i>	132717		Negative	
134	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	220523		<i>L. tropica</i> G1	162

135	<i>Ph. tobbi</i>	<i>Ph. tobbi</i>	66956		Negative	
136	<i>S. dentata</i>	<i>S. dentata</i>	174617		Negative	
137	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	148702		<i>L. donovani</i>	190
138	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	216154		Negative	
139	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	113247		Negative	
140	<i>Ph. tobbi</i>	<i>Ph. syriacus</i>	172020		Negative	
141	<i>S. fallax</i>	<i>S. dentata</i>	186429		Negative	
142	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	181734		<i>L. donovani</i>	150
143	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	187630		<i>L. tropica</i>	169
144	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	130176		Negative	
145	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	175317		Negative	
146	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	217771		<i>L. tropica</i> G1	754
147	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	219689		Negative	
148	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	189288		Negative	
149	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	444302		<i>L. tropica</i> G1	1532
150	<i>S. dentata</i>	<i>S. dentata</i>	323092		Negative	
151	<i>S. tiberiadis</i>	<i>S. ghesquierei</i>	192882		Negative	
152	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	477161		<i>L. tropica</i> G1	646
153	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	243772		Negative	
154	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	166009		Negative	
155	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	198128		Negative	
156	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	278152		Negative	
157	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	175154		Negative	
158	<i>S. fallax</i>	<i>S. dentata</i>	149616		Negative	
159	<i>S. tiberiadis</i>	<i>Ph. sergenti</i>	59830		Negative	
160	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	18137		Negative	
161	<i>S. fallax</i>	<i>S. dentata</i>	165508		Negative	
162	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	203261		Negative	

163	<i>S. schwetzi</i>	<i>S. schwetzi</i>	172715		Negative	
164	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	296688		Negative	
165	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	132403		Negative	
166	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	224322		Negative	
167	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	168474		Negative	
168	<i>Ph. sergenti</i>	<i>Ph. sergenti</i>	93608		Negative	
169	<i>S. dubia</i>	<i>S. dubia</i>	96612		Negative	
170	<i>S. dubia</i>	<i>S. dubia</i>	64419		Negative	
171	<i>Ph. papatasi</i>	<i>Ph. papatasi</i>	122724		Negative	
172		<i>Ph. sergenti</i> (Positive control)	102994		<i>L. tropica</i> G2	3347
173		<i>S. dentate</i> (Negative control)	1		Negative	

Appendix 5: work flow for NGS data analysis.



استعمال تقنية الجيل الثاني في كشف تسلسل الحمض النووي لمعرفة نوع ذبابة الرمل و طفيل الليشمانيا

إعداد: محمد هاشم حسين الطرده

المشرف الاول : د. زياد عابدين

المشرف الثاني: د. عبد المجيد ناصر الدين

5. الملخص

يعتبر داء الليشمانيات مرضاً شائعاً جداً في العالم، على اختلاف أنواعه، الباطنية والجلدية. الطفيل المسبب لداء الليشمانيات هو طفيل أحادي الخلية من نوع الليشمانيات يستطيع العيش والتكاثر في داخل خلايا الجهاز المناعي (Macrophages)

التابعة لجسم الانسان مضيف كما أن بإمكانه العيش والتكاثر، أيضاً، في أمعاء ذبابة الرمل، الناقل للمرض بين الحيوانات و الانسان.

دراسة هذا المرض و خصوصا الكائن الناقل لهذا المرض و هو ذبابة الرمل بعدة فصائلها مهم، لفهم كيفية انتقاله و الحد من الاصابة بهذا المرض. ان الطرق المتبعة في دراسة ذبابة الرمل تعتمد على الطرق التقليدية من خلال المجهر التشريحي و تمييز انواع ذبابة الرمل من خلال التشريح و ايجاد الفروقات بين فصائل الذباب وهذا يتطلب خبرة و جهد بالاضافة لوقت زمني اطول خصوصا عندما نتحدث عن عدد هائل من العينات.

قمنا في هذه الدراسة بالتقاط 171 عينة من ذبابة الرمل من منطقة طوباس شمال فلسطين، و تم عزل المادة الوراثية منها و بناء نظام يعتمد على تقنية الجيل الثاني في كشف تسلسل الحمض النووي NGS ومن خلال هذ النظام يمكننا التعرف على نوع الذبابة بالاضافة للتعرف على نوع طفيل الليشمانيا الذي كل ذبابة . اظهرت نتائج دراستنا الى وجود فصيل ذبابة الرمل *Phlebotomus* بنسبة (94.1%) و فصيل *Sergentomyia genera* بنسبة (5.9%) و الفصيل الاكثر تواجدا هو *Phlepetomus*

sergenti بنسبة (86%) . بالإضافة لوجود ذبابات الرمل الحاملة ل طفيل الليشمانيا المدارية *L. tropica* بنسبة (4.7%).

اظهرت نتائج دراستنا توافق مع نتائج الطريقة المجهرية التقليدية مع قيمة احتمالية $P < 0.001$. تم بناء هذا النظام كنظام بديل للطرق التقليدية, كما ان النظام الذي قمنا بتصميمه يتميز بدقة في تحديد فصيلة ذبابة الرمل وطفيل الليشمانيا في وقت اسرع من الطرق التقليدية.